Control of Hox transcription factor concentration and cell-to-cell variability by an auto-regulatory switch

Dimitrios K. Papadopoulos[1,*], Kassiani Skouloudaki[1], Ylva Engström[2], Lars Terenius[3], Rudolf Rigler[4,5], Christoph Zechner[1,6], Vladana Vukojević[3] and Pavel Tomancak[1,*]

[1]Max-Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany
[2]Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, 10691 Stockholm, Sweden
[3]Center for Molecular Medicine (CMM), Department of Clinical Neuroscience, Karolinska Institutet, 17176 Stockholm, Sweden
[4]Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 17177 Stockholm, Sweden
[5]Laboratory of Biomedical Optics, Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland
[6]Center for Systems Biology Dresden, 01307, Dresden, Germany
[*]Correspondence: papadopo@mpi-cbg.de and tomancak@mpi-cbg.de

Author contributions: DKP and VV, designed and performed the binding study of *Antp*, analyzed the data; DKP and KS, designed and performed the study of *Antp* relative transcript abundance, analyzed the data; YE, provided the *Antp* P1 and P2 reporter lines; DKP and PT, conceived the comparative analysis of TFs and the examination of Antp variability in second instar imaginal discs; CZ, developed the mathematical model of stochastic Antp expression, DKP, conceived the project, designed and performed all remaining experiments, analyzed the data, prepared figures and wrote the manuscript with critical input from KS, YE, CZ, VV and PT; all authors critically reviewed the manuscript and approved it for submission.

Abstract

Concentration of transcription factors (TFs) and their cell-to-cell protein variability are important functional determinants in development, yet how variability is controlled remains poorly understood. Using Fluorescence Correlation Spectroscopy (FCS), we characterized in live *Drosophila* imaginal discs the concentration and cell-to-cell variability of 14 endogenously tagged TFs. We found that the Hox TF Antennapedia (Antp) transitioned from a low concentration/high variability state early in development to a high concentration/low variability state later in development. Using FCS and temporally resolved genetic studies, we uncovered that Antp is necessary and sufficient to drive a developmental regulatory switch from auto-activation to auto-repression, thereby reducing variability. This switch is controlled by a progressive change in relative concentrations of preferentially activating and repressing Antp protein isoforms, which bind to chromatin with different affinities. We derived a simple mathematical model, confirming that the Antp auto-regulatory circuit would suffice to increase protein concentration while suppressing variability over time.

Introduction

In order to understand the mechanisms that control pattern formation and cell fate specification in developing organisms, the intranuclear concentration, DNA-binding kinetics and cell-to-cell variability of relevant TFs need to be quantitatively characterized. TF concentration variability at the tissue level is thought to arise from diverse processes, including mRNA transcription, protein production and degradation. For example, gene transcription in a given tissue is a noisy process. The noise is due to stochastic binding and interactions of proteins involved in

56 transcriptional activation of the specific gene (intrinsic noise) (Blake et al., 2003;
57 Elowitz et al., 2002) and also due to differences among cells in terms of abundance
58 of the transcriptional and post-transcriptional cellular machinery which affects the
59 efficiency of transcriptional activation in general (extrinsic noise) (Swain et al., 2002).
60 These can influence the expression and production of functional protein, resulting in
61 protein concentration that exhibits variability among cells.

62 In undifferentiated tissue or cells, TF cell-to-cell variability can be the driving
63 force for differentiation. For example, progressive establishment of a Nanog salt-and-
64 pepper expression pattern leads to the formation of primitive endoderm in the mouse
65 preimplantation embryo, whereas loss of the variability results in embryos lacking
66 primitive endoderm entirely (Kang et al., 2013). In Drosophila, the Senseless (Sens)
67 TF is required for the establishment of proper number of sensory organ precursors in
68 the ectodermal proneural cluster of cells and unequal concentration among cells is
69 required for their specification (Li et al., 2006). Moreover, variability in concentration
70 (rather than its overall average concentration) of the Yan TF drives the transition of
71 developing photoreceptor cells to a differentiated state during Drosophila eye
72 development (Pelaez et al., 2015).

73 Conversely, in already differentiated tissue or cells, TF expression variability
74 among cells may need to be counteracted to ensure homogeneity of gene
75 expression patterns and robustness of commitment to a certain transcriptional
76 regime. Such homogenization of expression levels has been identified for the Snail
77 (Sna) TF, which is required for the invagination of the mesoderm during Drosophila
78 gastrulation (Boettiger and Levine, 2013), or the Bicoid (Bcd) and Hunchback (Hb)
79 TFs during early embryogenesis (Gregor et al., 2007a; Gregor et al., 2007b; Little et
80 al., 2013). These studies have quantified the tolerable degrees of concentration

81 variability that allow establishment of gene expression territories with remarkable
82 precision in the developing embryo.

83 In addition, differential fates within the same developmental territory may be
84 specified by TFs deploying different DNA-binding dynamics despite the existence of
85 very similar concentrations (i.e. low variability). For example, studies on the Oct4 TF
86 in early mouse embryos have shown that differential kinetic behavior of DNA binding,
87 despite equal Oct4 concentration among blastomeres, ultimately dictates an early
88 developmental bias towards lineage segregation (Kaur et al., 2013; Plachta et al.,
89 2011).

90 So far, gene expression variability studies have focused predominantly on
91 monitoring the noise of mRNA production (Holloway et al., 2011; Holloway and
92 Spirov, 2015; Little et al., 2013; Lucas et al., 2013; Pare et al., 2009). Little
93 information exists about TF variability at the protein level within a tissue, since
94 studies of this sort would require the use of quantitative methods with single-
95 molecule sensitivity.

96 We have previously used Fluorescence Microscopy Imaging and FCS, to
97 study Hox TF interactions with nuclear DNA in living salivary gland cells
98 (Papadopoulos et al., 2015; Vukojevic et al., 2010). FCS has also been instrumental
99 for the quantification of TF dynamics in living cells or tissue in several recent studies
100 (Clark et al., 2016; Kaur et al., 2013; Lam et al., 2012; Mistri et al., 2015;
101 Papadopoulos et al., 2015; Perez-Camps et al., 2016; Szaloki et al., 2015; Tiwari et
102 al., 2013; Tsutsumi et al., 2016). Yet, in these studies, protein mobility has been
103 measured in overexpressing systems. However, to understand TF behavior *in vivo*,
104 proteins need to be quantified at endogenous levels (Lo et al., 2015).

105 In this study, we take advantage of the availability of fly toolkits, in which TFs
106 have been endogenously tagged by different methodologies: fosmid, BAC, FlyTrap
107 and MiMIC lines (Buszczak et al., 2007; Ejsmont et al., 2011; Ejsmont et al., 2009;
108 Kelso et al., 2004; Morin et al., 2001; Quinones-Coello et al., 2007; Sarov et al.,
109 2016; Venken et al., 2011), to measure the intranuclear concentration of various TFs
110 *in vivo* by FCS, and their cell-to-cell variability in several fly imaginal discs. The
111 imaginal discs are flat, single-layered epithelia comprised of small diploid cells and
112 many TFs are expressed in defined regions within these tissues during development.
113 In this system, we found that Antp, a well-characterized homeotic selector gene,
114 responsible for specification of thoracic identity in fly tissues, displayed high cell-to-
115 cell variability during early disc development and this variability was suppressed at
116 later developmental stages. Through a combination of genetics, single-molecule
117 measurements of TF dynamics by FCS and computational modeling, we uncovered
118 a mechanism that controls Antp concentration and variability in developing discs.
119

120 **Results**

121

122 Characterization of average protein concentrations and cell-to-cell variability of
123 *Drosophila* TFs

124 Average concentrations of TFs in neighboring nuclei of third instar imaginal
125 discs were measured by FCS (Figure 1 A-J and Figure 1 – figure supplement 1 A-P).
126 FCS is a non-invasive method with single molecule sensitivity, in which a confocal
127 arrangement of optical elements is used to generate a small (sub-femtoliter) volume
128 inside living cells, from which fluorescence is being detected (Figure 1 C and D,
129 green ellipsoid). Fluorescent molecules diffuse through this observation volume,

130 yielding fluorescence intensity fluctuations that are recorded over time by detectors
131 with single-photon sensitivity (Figure 1 E). These fluctuations are subsequently
132 subjected to temporal autocorrelation analysis, yielding temporal autocorrelation
133 curves (henceforth referred to as FCS curves, Figure 1 F), which are then fitted with
134 selected models to extract quantitative information about the dynamic processes
135 underlying the generation of the recorded fluctuations. In the case of molecular
136 movement of TFs (Supplement 1), information can be obtained regarding: a) the
137 absolute concentrations of TFs (Figure 1 F), (b) TF dynamic properties, such as:
138 diffusion times, differences in their interactions with chromatin and fractions of free-
139 diffusing *versus* chromatin-bound TFs (Figure 1 G); and c) TF cell-to-cell
140 concentration variability (Figure 1 H).

141 For the 14 selected TFs, we measured average concentrations ranging about
142 two orders of magnitude among different TFs, from $\sim 30\ nM$ to $\sim 1.1\ \mu M$ (Figure 1 I
143 and Figure 1 – figure supplement 1 A-Q). We also obtained various diffusion times
144 and fractions of slow and fast diffusing TF molecules (Figure 1 J), indicating
145 differential mobility and degree of DNA-binding among different TFs (Vukojevic et al.,
146 2010). Comparison of the y-axis amplitudes at the smallest lag time of the FCS
147 curves (the points at which the FCS curves cross the y-axis), which are inversely
148 proportional to the concentration of fluorescent molecules (Figure 1 F), gives
149 information about concentration variability (heterogeneity) among different cell
150 nuclei, i.e. reflects heterogeneity of protein concentration at the tissue level (Figure 1
151 H). We measured the variability of all 14 TFs in our dataset (expressed as the
152 variance over the mean squared, $CV^2 = \frac{s^2}{m^2}$) to be in the range $7 - 37\%$ (Figure 1 K
153 and Figure 1 – figure supplement 1 Q). These numbers are consistent with previous

154 observations (Sanchez et al., 2011). We then used this dataset as a starting point for
155 studying the control of variability during imaginal disc development.

156 Since low variability at the tissue level is likely to be achieved through some
157 active mechanism that controls it, we searched for TFs that exhibited low variability
158 and relatively high concentrations. One TF, the Hox gene Antp, had comparatively
159 lower variability ($CV^2 < 0.2$) for its high average concentrations (Figure 1 K), in
160 particular in the leg disc. This distinction prompted us to measure variability of Antp
161 at different concentrations by examining clusters of neighboring cells from across the
162 disc displaying different average expression levels (Figure 1 L). Because FCS
163 performs best at low to moderate expression levels, we performed this analysis in
164 the wing disc (Figure 1 L). We established that the observed fluorescence intensity
165 fluctuations were caused by diffusion of TF molecules through the illuminated
166 confocal volume (Figure 1 – figure supplement 2 and Supplement 1). Our data
167 showed that Antp cell-to-cell variability decreased with increasing Antp concentration
168 (Figure 1 M), suggesting complex transcriptional regulatory processes (Franz et al.,
169 2011; Smolander et al., 2011) that we further investigated using the powerful
170 *Drosophila* genetic toolkit.

171

172 Control of Antp concentration by transcriptional auto-regulation

173 One mechanism by which genes control their expression level variability is
174 auto-regulation (Becskei and Serrano, 2000; Dublanche et al., 2006; Gronlund et al.,
175 2013; Nevozhay et al., 2009; Shimoga et al., 2013; Thattai and van Oudenaarden,
176 2001). To test whether Antp can regulate its own protein levels, we monitored the
177 concentration of endogenous Antp protein upon overexpression of *Antp* from a
178 transgene. To distinguish between overexpressed and endogenous protein, we used

179 synthetic Antp (SynthAntp) transgenes fused to eGFP (SynthAntp-eGFP). These
180 transgenes encode the Antp homeodomain, the conserved YPWM motif and the C
181 terminus (but lack the long and non-conserved N terminus of the protein, against
182 which Antp antibodies have been raised) and they harbor Antp-specific homeotic
183 function (Papadopoulos et al., 2011). Clonal overexpression of *SynthAntp-eGFP* in
184 the wing disc notum (Figure 2 A-B' and controls in Figure 2 – figure supplement 1 D
185 and D') repressed the endogenous Antp protein, indicating that Antp is indeed able
186 to regulate its own protein levels.

187 Since Antp is a TF, we next asked whether the auto-repression indeed occurs
188 at the transcriptional level. The *Antp* locus is subject to complex transcriptional
189 regulation, involving a distal and a proximal promoter (P1 and P2 promoters,
190 respectively), spanning more than 100 kb of regulatory sequences. We established
191 that the P1 promoter (rather than the P2 promoter) is predominantly required to drive
192 expression of Antp in the wing disc notum (Figure 2 – figure supplement 1 A-C'), in
193 line with previous observations ((Engstrom et al., 1992; Jorgensen and Garber,
194 1987; Zink et al., 1991) and Materials and Methods). Moreover, mitotic
195 recombination experiments in regions of the wing disc unique to P2 transcription
196 have shown no function of the P2 promoter transcripts in wing disc development
197 (Abbott and Kaufman, 1986). Thus, the P1 Antp reporter serves as a suitable
198 reporter of the *Antp* locus transcriptional activity in this context.

199 Clonal overexpression of SynthAntp-eGFP in the wing disc repressed the
200 *Antp* P1 transcriptional reporter (Figure 2 C and C' and controls in Figure 2 – figure
201 supplement 1 E and E'). To rule out putative dominant negative activity of the small
202 synthetic Antp-eGFP peptide, we also performed these experiments with the full-
203 length Antp protein (Figure 2 – figure supplement 1 F and F'). We conclude that the

204 Antp protein is able to repress its own transcription from the P1 promoter, suggesting
205 a possible mechanism of suppressing cell-to-cell variability of Antp expression levels
206 (Figure 2 D).

207 In the course of these experiments, we noticed that ectopic overexpression of
208 *SynthAntp-eGFP* or the full-length Antp protein from the *Distal-less* (*Dll*) (MD23)
209 enhancer surprisingly resulted in activation of the *Antp* P1 reporter in distal
210 compartments of the wing disc, such as the wing pouch, where Antp is normally not
211 detected (Figure 2 E-F' and controls in figure 2 – figure supplement 1 G-H'). This
212 finding suggests that next to its auto-repressing function, Antp is also capable to
213 activate its own transcription (Figure 2 G).

214 To exclude that the auto-activation and repression of Antp are artifacts of
215 overexpression, we measured by FCS the concentration of Antp triggered by
216 different Gal4-drivers and found it to directly correlate with the degree of its homeotic
217 transformation capacity (Figure 2 – figure supplement 2 A-I). Importantly, we
218 observed indistinguishable DNA-binding behavior not only across the whole
219 concentration range examined (Figure 2 – figure supplement 2 J), but also between
220 endogenous and overexpressed Antp (Figure 2 – figure supplement 3 A-B).
221 Importantly, the auto-activating and auto-repressing capacity of Antp was preserved
222 even with the weak Gal4-driver *69B* (Figure 2 – figure supplement 2 K-P) that
223 triggered concentrations of Antp slightly lower than its normal concentration in the
224 wing disc ($473\,nM$ versus $501\,nM$), indicating that auto-activation and auto-
225 repression of Antp take place at endogenous concentrations.

226 We conclude that, Antp is able to repress and activate its own transcription
227 (Figure 2 D and G) and hypothesize that this auto-regulatory circuit somehow sets
228 the correct concentration of Antp protein in imaginal discs.

229

230 A temporal switch controls the transition of *Antp* from a state of auto-activation to a
231 state of auto-repression

232 To further investigate the mechanism by which the Antp auto-regulatory circuit
233 sets the Antp expression levels precisely, we next asked whether the seemingly
234 opposing auto-regulatory activities of Antp are separated in time during
235 development. To that end, we induced gain-of-function clones of full-length untagged
236 *Antp* either at 26 h (first larval instar – henceforth referred to as "early" stage) or at
237 60 h (late second larval instar – henceforth referred to as "late" stage) of
238 development and analyzed the clones in late third instar wing imaginal discs (Figure
239 3). As a pre-requisite for this analysis, we established that the Antp-eGFP
240 homozygous viable MiMIC allele recapitulates the endogenous Antp pattern in the
241 embryo and all thoracic imaginal discs and therefore can be used to monitor
242 endogenous Antp protein (Figure 3 – figure supplement 1). Clonal induction of full-
243 length untagged *Antp* in early development triggered strong auto-activation of Antp-
244 eGFP (Figure 3 A, B and B' and controls in Figure 3 – figure supplement 2 A-C'). As
245 before, we confirmed that early auto-activation of Antp is transcriptional and similar
246 for both full-length and SynthAntp proteins (Figure 3 – figure supplement 2 D-E' and
247 controls in F-G'). Early auto-activation was further supported by a loss-of-function
248 experiment, where *RNAi*-mediated early knockdown of Antp resulted in
249 downregulation of the Antp reporter (Figure 3 C and C' and controls in Figure 3 –
250 figure supplement 2 H and H'). The loss and gain-of-function analysis together
251 suggest that during early disc development Antp is required for sustaining its own
252 expression.

253 In contrast, clonal induction during the late second instar stage (Figure 3 F)
254 repressed Antp-eGFP (Figure 3 G and G') and, reciprocally, the clonal knockdown
255 by *RNAi* triggered auto-activation of Antp transcription (Figure 3 H and H'). Hence, in
256 contrast to early development, Antp represses its own expression in third instar
257 discs.

258 While the gain-of-function experiments show that Antp is sufficient to execute
259 auto-regulation, loss-of-function analysis indicates that it is also necessary for both
260 repression and activation at the transcriptional level.

261 Together, these results revealed the existence of a switch in Antp auto-
262 regulatory capacity on its own transcription during development. Starting from a
263 preferentially auto-activating state early in development (Figure 3 D), it changes into
264 an auto-inhibitory mode at later developmental stages (Figure 3 I).

265

266 <u>Antp switches from a low-concentration/high-variability to a high-concentration/low-</u>
267 <u>variability state</u>

268 If the *Antp* auto-repressive state limits the variability of Antp protein
269 concentration among neighboring cells late in development, we expected that the
270 variability would be higher during earlier stages, when auto-repression does not
271 operate. We, therefore, monitored the endogenous expression levels and cell-to-cell
272 variability of Antp nuclear concentration in second instar wing and leg discs by FCS.
273 We observed significantly lower average concentrations of Antp protein in second
274 versus third instar wing and leg discs and the inverse was true for concentration
275 variability (Figure 3 E and Figure 3 – figure supplement 3 A, A' and C), indicating that
276 the developmental increase in concentration is accompanied by suppression of
277 concentration variability. In addition, FCS analysis revealed a notable change in Antp

278 characteristic decay times (signifying molecular diffusion, limited by chromatin-
279 binding) at early versus late stages (Figure 3 E and Figure 3 – figure supplement 3
280 B). This behavior indicates that endogenous Antp is initially fast moving in the
281 nucleus and undergoes considerably fewer interactions with chromatin, compared to
282 later stages where its interactions with chromatin are more frequent and longer
283 lasting.

284 Taken together, our FCS measurements show that *Antp* is expressed at
285 relatively low and highly variable levels in early developing discs, when genetic
286 evidence indicates auto-activation capacity on its own transcription. Later in
287 development, when Antp has reached a state of higher average concentrations,
288 auto-repression kicks in, resulting in considerably lower variability among
289 neighboring cells.

290

291 Dynamic control of Antp auto-regulation by different Antp isoforms

292 The changing binding behavior of Antp on chromatin from second to third
293 instar discs and the developmental transition from an auto-activating to an auto-
294 repressing state suggested a causal relationship between the two phenomena. We,
295 therefore, sought to identify molecular mechanisms that could link the observed
296 changes in Antp chromatin-binding to Antp auto-activation and repression. It is well
297 established that the Antp mRNA contains an alternative splice site in exon 7
298 immediately upstream of the homeobox-containing exon 8, and generates Antp
299 isoforms differing in as little as 4 amino acids in the linker between the YPWM motif
300 (a cofactor-interacting motif) and the homeodomain (Figure 4 A) (Stroeher et al.,
301 1988). Our previous observation that long linker isoforms favor transcriptional
302 activation of Antp target genes, whereas short linker ones favor repression of Antp

303 targets (Papadopoulos et al., 2011), prompted us to examine whether the linker
304 length is also responsible for differences in auto-regulation.

305 Ectopic expression of SynthAntp-eGFP peptides featuring a long linker
306 displayed significantly weaker repression capacity on endogenous Antp, as
307 compared to their short linker counterparts (Figure 4 B, B', F and F' and quantified in
308 D and H, see also Materials and Methods). We confirmed that, also in this case, the
309 repression was at the transcriptional level (Figure 4 – figure supplement 1 I-J').
310 Inversely, long linker Antp isoforms exhibited stronger activation of Antp reporter, as
311 compared to short linker ones (Figure 4 C, C', G, G' and quantified in D and H, see
312 also Materials and Methods). We, additionally, validated that short linker isoforms
313 encoded by full-length or SynthAntp cDNAs behaved as weaker auto-activating and
314 stronger auto-repressing Antp species in all our previous experiments of
315 endogenous Antp protein and P1 reporter (Figure 4 – figure supplement 1 A-H'). We
316 conclude that, also in the case of Antp auto-regulation, short linker isoforms function
317 as more potent repressors, whereas long linker ones operate as more potent
318 activators.

319 Since the Antp P1 promoter unit changes its configuration from preferential
320 auto-activation to auto-repression, and short and long linker Antp isoforms function
321 as preferential auto-repressors and auto-activators, it appeared possible that the
322 switch in Antp regulation is executed at the level of transcript variant abundance of
323 these isoforms. Therefore, we next quantified the relative abundance of long and
324 short linker transcript variants in the embryo, second and third instar discs (Figure 4
325 D and H). We found that the concentration of the long linker variant decreased,
326 whereas the concentration of the short linker variant increased over time in
327 development, in line with previous observations (Stroeher et al., 1988). As

328 hypothesized, this finding suggested that relative transcript variant abundance may
329 underlie the switch between auto-activation and auto-repression (without excluding
330 additional mechanisms).

331 Relative changes in Antp transcript variant concentration (Figure 4 D and H),
332 differential efficiency of their encoding isoforms to repress or activate the *Antp* gene
333 (Figure 4 B-D and F-H), the developmental switch of the *Antp* gene from auto-
334 activation to repression (Figure 3) and the different mobilities of Antp between
335 second and third instar imaginal discs (Figure 3 E) all pointed towards the hypothesis
336 that the two isoforms have different properties in their modes of interaction with
337 chromatin. To investigate this, we expressed the two isoforms in third instar wing and
338 antennal discs from the *69B* enhancer, which we established to result in Antp
339 concentrations close to (if not below) endogenous levels (Figure 2 – figure
340 supplement 2 A-J). FCS measurements revealed that the short linker isoform
341 displayed longer characteristic decay times and higher fraction of DNA-bound
342 molecules, suggesting stronger and more pronounced binding to chromatin than its
343 long linker counterpart (Figure 4 D and H and Figure 4 – figure supplement 2 A-B).
344 With chromatin (and therefore Antp binding sites configuration) being identical
345 between the two instances (short and long linker isoforms examined in third instar
346 wing and antennal imaginal discs of the same age), we were able to directly
347 compare the apparent equilibrium dissociation constants for the two isoforms
348 (Supplement 3). We found that the affinity of binding to chromatin ($K_d{}^{-1}$) of the
349 repressing short linker isoform is at least 2.3 times higher compared to the activating
350 long linker isoform ($\frac{K_{d,Antp}^{long\ linker\ isof.}}{K_{d,Antp}^{short\ linker\ isof.}} > 2.3$) (Figure 4 D and H and Figure 4 – figure
351 supplement 2 C-D'). We, additionally, validated the different affinities of short and
352 long isoforms by gel-shift experiments using different Antp binding sequences and

353 obtained from two- to eightfold higher affinity of the short linker isoform compared to
354 the long linker isoform in binding of previously characterized Antp and homeodomain
355 binding sites (Figure 4 D and H and Figure 4 – figure supplement 3). Collectively, our
356 experiments support the notion that differences in Antp regulation during disc
357 development can be largely attributed to differences in the affinity of the investigated
358 Antp isoforms.

359 Taken together, the switch of Antp from an auto-activating to an auto-
360 repressing state and the alteration of its DNA-binding behavior during disc
361 development can be largely explained by a temporal developmental regulation of the
362 relative concentrations of preferentially auto-activating and auto-repressing Antp
363 protein isoforms, which themselves display distinct properties in their modes of
364 interaction with chromatin (Figure 4 E and I).

365

366 Robustness of Antp auto-regulation

367 In order to further substantiate the qualitative model of Antp auto-regulation
368 suggested by our experimental findings, we developed a mathematical model of
369 stochastic *Antp* expression. This model tests whether the identified interplay
370 between positive and negative auto-regulation of Antp through distinct isoforms is
371 sufficient to explain the increase in protein concentration and decrease in nucleus-to-
372 nucleus variability from early to late stages. The model consists of a dynamic
373 promoter, which drives transcription of *Antp* followed by a splicing step, leading to
374 the expression of either the auto-repressing or the auto-activating isoform of Antp. In
375 line with our finding that the repressing isoform has higher concentration at later
376 stages, we assumed that splicing is more likely to generate this isoform than the
377 activating isoform. The initial imbalance of Antp towards the activating isoform

378 (Figure 4 D and H) is modeled through appropriate initial concentrations of each
379 isoform.

380 Since Antp copy numbers per nucleus are in the thousands at both early and
381 late stages, intrinsic noise of gene expression is likely to explain only a certain
382 portion of the overall variability in Antp concentrations (Elowitz et al., 2002; Taniguchi
383 et al., 2010). The remaining part (termed extrinsic variability) is due to cell-to-cell
384 differences in certain factors affecting gene expression such as the ribosomal or ATP
385 abundances. To check whether extrinsic variability significantly affects Antp
386 expression, we expressed nuclear RFP constitutively, alongside with endogenous
387 Antp and measured the abundances of green-labeled Antp and RFP. Since extrinsic
388 factors are expected to affect both genes in a similar way, one should observe a
389 correlation between the concentration of nuclear RFP and Antp-eGFP. Our data
390 showed a statistically significant correlation between RFP and Antp (Figure 5 – figure
391 supplement 1, $r = 0.524$ and $p = 9.77 \cdot 10^{-5}$). Correspondingly, we accounted for
392 extrinsic variability also in our model by allowing gene expression rates to randomly
393 vary between cells (Zechner et al., 2012).

394 The promoter itself is modeled as a Markov chain with three distinct
395 transcriptional states. In the absence of Antp, the promoter is inactive and
396 transcription cannot take place (state “U” in Figure 5 A). From there, the promoter
397 can switch into a highly expressing state “A” at a rate that is assumed to be
398 proportional to the concentration of the long-linker, auto-activating isoform. This
399 resembles the positive auto-regulatory function of Antp. Conversely, the promoter
400 can be repressed by recruitment of the short-linker, auto-repressing isoform,
401 corresponding to state “R” in the model (Figure 5 A). To account for potential
402 leakiness of the promoter, this rate is not assumed to be zero, but significantly lower

403 than that of state “A”. Since the auto-repressing isoform of Antp can also activate the
404 promoter, albeit significantly weaker than the auto-activating isoform, and vice versa,
405 we allow the promoter to switch between states “A” and “R”.

406 While this promoter model resembles the dual-feedback structure of *Antp*
407 locus inferred from experiments, it is unclear whether the two isoforms compete for
408 the same binding sites on the P1 promoter or if auto-repression can take place
409 regardless of whether an activating isoform is already bound to the promoter. In the
410 former case, an increase in concentration of repressing Antp species enhances the
411 probability to reach state “R” only if the promoter is in state “U” (Figure 5 B and B’’).
412 In the latter case, also the rate of switching between “A” and “R” depends on the
413 concentration of repressing isoforms of Antp (Figure 5 A-A’’). We analyzed both
414 model variants by forward simulation and found that both of them can explain the
415 increase in average Antp concentration between early and late stages (Figure 5 A’’
416 and B’’), as well as the relative fraction of repressing and activating isoforms (Figure
417 5 D and D’). However, only the non-competitive binding model (Figure 5 A) can
418 explain the substantial reduction of total Antp variability between early and late
419 stages (Figure 5 A’), whereas in the competitive model variability is not reduced
420 (Figure 5 B’). We additionally established that the negative feedback is required for
421 suppression of variability (Figure 5 C-C’’), since without this, no suppression of
422 variability is conferred (Figure 5 C’). Thus, our model suggested that auto-repression
423 is required and it is possible also if an activating isoform of Antp is already bound to
424 the P1 promoter. Correspondingly, we use the non-competitive promoter model for
425 further analyses.

426 To further validate our model, we first examined its predictions on variability
427 by comparing the variability values predicted by the model to the ones generated by

428 our experimental measurements. Next to the CV^2, the Fano factor (expressed as the
429 variance over the mean, $FF = \frac{s^2}{m}$, in concentration units) is another commonly used
430 index to quantify variability in biological systems. Fano factor values that increase
431 with average concentrations indicate that the underlying transcriptional processes
432 cannot be sufficiently explained by a simple one-step promoter configuration with
433 purely intrinsic Poissonian noise and that extrinsic noise is likely to contribute
434 significantly to the overall variability (Newman et al., 2006; Schwanhausser et al.,
435 2011; Taniguchi et al., 2010). Our model predicted a decrease in variability as a
436 function of total Antp concentration and an increase in the Fano factor. These
437 findings are in good agreement with our experimental data (compare Figure 5 E to E'
438 and F to F').

439 We next analyzed the model behavior under different genetic perturbations.
440 We found that overexpression of both auto-activating and auto-repressing isoforms
441 leads to an increase of the total Antp concentration (Figure 5 G' and H'), but there is
442 no negative effect on the noise suppressing property of the circuit (Figure 5 G and
443 H). In fact, the variability is even further decreased, which can be explained by the
444 characteristic inverse relation between intrinsic noise and average concentration
445 (Paulsson, 2004). In line with this prediction, flies expressing roughly eightfold
446 concentration of either SynthAntp auto-activating or auto-repressing isoform in distal
447 appendages (Figure 5 I and J, Figure 2 – figure supplement 2 A and I) or the notum
448 (Figure 5 K and Figure 5 – figure supplement 2 A and A') displayed the wild type
449 morphology, indicative of normal Antp function.

450 However, overexpression of an exogenous repressor, such as Sex combs
451 reduced (Scr), which can only repress Antp at the transcriptional level, but can
452 neither activate it nor activate its own transcription (Figure 5 – figure supplement 2 B-

453 G'), was predicted by the model to block transcription almost entirely and to have
454 correspondingly severe effects on Antp dynamics (Figure 5 – figure L and L'),
455 leading to high variability even at late stages and to very low expression levels. As
456 expected, flies overexpressing SynthScr in the distal appendages or the notum,
457 displayed severe malformations of both leg and notum development (Figure 5 M-O).
458 Taken together, the minimal model of Antp auto-regulatory genetic circuit is
459 able to explain the experimentally observed differences in Antp concentration and
460 cell-to-cell variability at early and late developmental stages.
461

462 ## Discussion

463 In this work, we have characterized the endogenous molecular numbers
464 (concentration) and cell-to-cell variability in concentration of 14 TFs in *Drosophila*
465 imaginal discs by FCS. We have identified Antp as a TF displaying, for its high
466 average concentrations, considerably lower variability among cells. We used a
467 combination of genetics, FCS and mathematical modeling to quantitatively
468 characterize Antp behavior in live imaginal discs and identified a kinetic mechanism
469 responsible for the suppression of variability in third instar discs compared to earlier
470 developmental stages. We found that *Antp* can auto-regulate its expression levels
471 during the course of development, starting from a preferentially auto-activating state
472 early in development and transitioning to a preferentially auto-repressing state later.
473 The early state is characterized by lower average Antp concentrations and high
474 variability, whereas the opposite is true for the later repressing state. Without
475 excluding other mechanisms, such as chromatin configuration and accessibility of
476 Hox binding sites to Antp, we showed that differential expression of Antp isoforms is
477 one contributing mechanism for the observed regulatory switch. These isoforms

478 have preferentially activating or repressing activities on the Antp promoter, bind
479 chromatin with different affinities and are themselves expressed in different relative
480 amounts during development. Finally, based on this data, we have derived a simple
481 kinetic model of *Antp* auto-regulation and confirmed its predictions by introducing
482 genetic perturbations.

483 Negative auto-regulation has been identified as a frequently deployed
484 mechanism for the reduction of noise (cell-to-cell variability) and the increase of
485 regulatory robustness in various systems (Becskei and Serrano, 2000; Dublanche et
486 al., 2006; Gronlund et al., 2013; Nevozhay et al., 2009; Shimoga et al., 2013; Thattai
487 and van Oudenaarden, 2001). Auto-repression has been described for the Hox gene
488 *Ultrabithorax* (*Ubx*) in haltere specification and as a mechanism of controlling Ubx
489 levels against genetic variation (Crickmore et al., 2009; Garaulet et al., 2008), as well
490 as in *Ubx* promoter regulation in Drosophila S2 cells (Krasnow et al., 1989). In
491 contrast, an auto-activating mechanism is responsible for the maintenance of
492 *Deformed* expression in the embryo (Kuziora and McGinnis, 1988). Moreover, global
493 auto-regulation of *Hox* gene complexes has been shown to be in effect also in
494 mammalian limb development (Sheth et al., 2014). These experiments point to
495 evolutionarily conserved mechanisms for establishing (auto-activation) or limiting
496 (auto-repression) Hox TF levels and variability in different developmental contexts.

497 Our data suggest that the developmental switch from auto-activation to auto-
498 repression is, at least in part, mediated by molecularly distinct Antp linker isoforms.
499 Differences in affinities of different Hox TF isoforms, based on their linker between
500 the YPWM motif and the homeodomain, have also been identified for the Hox TF
501 Ubx. Interestingly, its linker is also subject to alternative splicing at the RNA level
502 (Reed et al., 2010). In a similar way to Antp, the long linker Ubx isoform displays 4-

503 5fold lower affinity of DNA binding, as compared to short linker isoforms, and the two
504 isoforms are not functionally interchangeable in *in vivo* assays. Finally, the Ubx linker
505 also affects the strength of its interaction with the Hox cofactor Extradenticle (Exd),
506 underscoring the functional importance of linker length in Hox TF function (Saadaoui
507 et al., 2011). Thus, protein isoform control might represent a common regulatory
508 mechanism of Hox-mediated transcriptional regulation.

509 Our model predicted that the Antp auto-regulatory circuit is robust with respect
510 to initial conditions and extrinsic noise by being able to suppress cell-to-cell
511 concentration variability even at very high concentrations of the auto-repressing or
512 the auto-activating Antp isoform. This "buffering" capacity on cell-to-cell variability is
513 reflected in the ability of flies to tolerate up to 7-fold overexpression of Antp without
514 exhibiting abnormal phenotypes. Therefore, two different isoforms produced from the
515 same gene with opposing roles in transcriptional regulation and different auto-
516 regulatory binding sites on the gene's promoter seem to suffice to create a robust
517 gene expression circuit that is able to "buffer" perturbations of the starting conditions.
518 So far, we have only been able to indiscriminately increase or decrease Antp
519 concentration at the tissue level and record the phenotypic outcome of these
520 perturbations. It will be interesting to test whether controlled perturbations of TF
521 variability at the tissue level (making TF concentration patterns less or more noisy
522 among neighboring cells) lead to abnormal phenotypes, however the technology for
523 such manipulation is currently not readily accessible in flies.

524 While our study has focused on the quantitative analysis of the influence of
525 Antp on its own expression, it may also have important implications for the regulation
526 of Antp target genes. In particular, the repression and activation of genes through
527 different isoforms of the same TF represents a plausible design principle to achieve

528 differential expression of target genes. This could be achieved either through
529 temporal developmental control of isoform abundance, or by spatial control, for
530 example through the formation of different nuclear microenvironments.

531 In the case of temporal control, since target genes harbor different amounts of
532 binding sites for activating and repressing isoforms, transcriptional programs of cells
533 could be easily switched by changes in relative concentration of TF isoforms. In the
534 case of Antp, for instance, targets that allow binding of only the short-linker,
535 repressing isoform, could be highly expressed initially, but expression would be shut
536 down at later stages as soon as this isoform becomes dominant. Conversely, the
537 opposite behavior would be observed for targets featuring binding sites for only the
538 activating isoform. We previously established that Antp target genes are activated or
539 repressed with different efficiencies by Antp isoforms (Papadopoulos et al., 2011).
540 Such a mechanism would allow distinct sets of target genes to be differentially
541 regulated by the same Hox TF at different developmental stages. Evidence from
542 genome-wide investigation of Hox target sites in the wing disc during development
543 (third instar larval, prepupal and pupal wing discs) has shown that target gene
544 batteries change dramatically in development (Pavlopoulos and Akam, 2011).

545 In the case of spatial control, enhancer binding sites of similar affinity could
546 cluster in topologically distinct regions in the nucleus, according to the highest
547 concentrations of either the preferentially activating or the repressing Antp isoform,
548 thus creating microenvironments, capable of favoring transcriptional activation or
549 repression. In this case, developmental control of the abundances of these two
550 isoforms would change the efficiencies of activation and repression of their targets
551 by enrichment or shrinkage of the microenvironments and/or departure/arrival of
552 enhancers to pre-existing microenvironments. We have already observed that Antp

553 nuclear distribution is not even, but features sites of accumulation (Figure 5 – figure
554 supplement 1 B). While this scenario remains to be examined in detail for Antp,
555 recent work on the Ubx-dependent expression of the *shavenbaby* (*svb*) gene
556 identified that this mechanism of generation of functionally distinct nuclear
557 microenvironments not only exists, but also allows robust expression of the *svb* locus
558 from low-affinity enhancers (Crocker et al., 2015; Crocker et al., 2016; Crocker et al.,
559 2017).

560 In general, while this work has increased our understanding of how
561 developmentally important TFs secure their own regulatory robustness, it will be
562 interesting to investigate whether the design principles of auto-regulatory circuits
563 extend also to target genes.

564

565 ## Figure legends

566 Figure 1: **Study of concentration, DNA-binding dynamics and cell-to-cell**
567 **protein concentration variability of 14 Drosophila TFs in imaginal discs.** (A-H)
568 Workflow of the study of TF concentration, mobility/DNA-binding and nucleus-to-
569 nucleus variability of TFs in imaginal discs by FCS. Live imaging of imaginal discs,
570 expressing endogenously-tagged TFs, visualized by fluorescence microscopy and
571 neighboring cells, expressing TFs at different levels, selected for FCS
572 measurements (A-B). FCS measurements are performed by placing the focal point
573 of the laser light into the nucleus (C-D) and recording fluorescence intensity
574 fluctuations (E), generated by the increase or decrease of the fluorescence intensity,
575 caused by the arrival or departure of fast- and slowly-diffusing TF molecules into or
576 out of the confocal detection volume (D). The recorded fluctuations are subjected to
577 temporal autocorrelation analysis, which generates temporal autocorrelation curves

578 (henceforth referred to as FCS curves), which by fitting with an appropriate model
579 (Supplement 1), yield information about the absolute concentration of fluorescent
580 molecules (F) and, after normalization to the same amplitude, their corresponding
581 diffusion times, as well as the fraction of fast- and slowly-diffusing TF molecules (G).
582 The concentration of molecules is inversely proportional to the y-axis amplitude at
583 the origin of the FCS curve (F). Processes that slow down the diffusion of TF
584 molecules, such as binding to very large molecules (e.g. chromosomal DNA), are
585 visible by a shift of the FCS curves to longer characteristic times (G). Measurements
586 in a collection of neighboring cell nuclei also allow the calculation of protein
587 concentration variability at the live tissue level (H). (I) Representative average FCS
588 measurements of eight TFs, selected to demonstrate the full span of the
589 concentration range observed. (J) FCS curves shown in (I), normalized to the same
590 amplitude, $G_n(\tau) = 1$ at $\tau = 10\ \mu s$, show the shift of the characteristic decay times to
591 higher values and the increase of the relative amplitude of the slow component,
592 indicating slower diffusion and/or higher degree of TF binding to the DNA. (K)
593 Variability of the 14 TFs as a function of concentration. Antp in the wing, but even
594 more in the leg, disc displays markedly lower variability for its average
595 concentrations. (L) Variability in concentration of endogenous Antp in the wing disc,
596 observed among clusters of neighboring cells, each displaying a different average
597 concentration. (M) Variability of Antp concentration in clusters of neighboring cell
598 nuclei (plotted as a function of its average concentrations), decreases as the
599 average Antp concentration in the cluster increases, i.e. low average concentration
600 clusters are associated with high variability, whereas high average concentration
601 clusters exhibit low variability. Error bars in (K) and (N) represent 1 standard
602 deviation of average TF concentration.

603

604 Figure 1 – figure supplement 1: **Measurement of average concentrations and**
605 **nucleus-to-nucleus variability of 14 endogenously-tagged TFs in Drosophila**
606 **imaginal discs by FCS.** (A-P) Fluorescence imaging of TFs, showing their
607 expression pattern in imaginal discs and the salivary gland. White arrows indicate
608 regions where FCS measurements of endogenous intra-nuclear concentration were
609 performed and the average concentrations are given for each TF. Images have been
610 contrasted for visualization purposes. For the Antp and Grn TFs, both leg and wing
611 imaginal discs have been used for measurements. Average concentrations of TFs
612 measured in different cells span a range of two orders of magnitude, from few tens to
613 a thousand nanomolar. Scale bars denote $100\,\mu m$, unless otherwise indicated. (Q)
614 Characterization of nucleus-to-nucleus variability among neighboring cells within the
615 same expression domain in imaginal discs of the 14 TF studied by FCS. Black bars
616 show concentration averages (with error bars representing 1 standard deviation),
617 whereas grey bars show the variability, i.e. the squared coefficient of variability
618 (expressed as the variance over the squared mean, $CV^2 = \frac{s^2}{m^2}$). TFs have been
619 sorted according to increasing variability.

620

621 Figure 1 – figure supplement 2: **Characteristic decay times of Antp-eGFP do not**
622 **change as a function of total concentration.** (A-B) Characteristic decay times τ_{D_1}
623 (A) and τ_{D_2} (B) do not vary with the concentration of Antp-eGFP TF molecules, as
624 evident from $\tau_{D_1} = f\,(N_1)$ and $\tau_{D_2} = f\,(N_2)$, where N_1 is the number of freely
625 diffusing, N_2 the number of bound Antp-eGFP TF molecules and τ_{D_1}, τ_{D_2} their
626 respective diffusion times.

627

628 Figure 2: **Antp is able to repress and activate its own expression at the**
629 **transcriptional level.** (A) Schematic representation of the wing disc showing the
630 region of highest Antp expression (green cells). Antp is highly expressed in the wing
631 disc in the regions of the notum that correspond to the structure of the prescutum in
632 the adult cuticle, as well as in the base of the wing blade, which gives rise to the
633 mesopleura and pteropleura of the adult thoracic cuticle. The black rectangle
634 indicates the approximate region where the clonal analysis has been performed in
635 (B-C'). (B-B') Clonal overexpression of a *SynthAntp-eGFP* construct (green) in the
636 wing disc represses the endogenous Antp protein (magenta). The region enclosed
637 by a dashed line in (B) shows a clone induced in the notum within the *Antp*
638 expression domain. (C-C') Auto-repression of Antp occurs at the transcriptional level.
639 Induction of *SynthAntp-eGFP*-overexpressing clones, while monitoring *Antp*
640 transcription by the *Antp* P1-*lacZ* reporter. Dashed lines indicate clones within the
641 *Antp* P1 reporter expression domain that express *SynthAntp-eGFP* and down-
642 regulate *Antp* P1-*lacZ*. (D) Schematic representation of Antp repressive function on
643 its own transcription. (E) Schematic representation of the wing disc. The black
644 rectangle indicates the approximate region of ectopic Antp P1 reporter expression in
645 (F-F'). (F-F') Expression of *SynthAntp-eGFP* by *Dll*-Gal4 (MD23) results in ectopic
646 activation of *Antp* P1-*lacZ* in distal compartments of the wing disc. Note the
647 endogenous and ectopic expression pattern of the reporter in (F), indicated by white
648 arrows. (G) Schematic representation of Antp activating function on its own
649 transcription. Scale bars denote $100\ \mu m$.
650
651 Figure 2 – figure supplement 1: **Antp is able to repress and activate itself at the**
652 **transcriptional level – controls.** (A-C') Normal expression patterns of the Antp P1

653 (A-A') and P2 (B-B') transcriptional reporters and Antp protein immunohistochemistry
654 (C-C'). Boxed areas in (A), (B) and (C) are magnified in (A'), (B') and (C'). The Antp
655 P1 reporter is highly expressed in the prescutum region of the notum (A') and the
656 peripodial cells at the base of the wing blade (giving rise to the mesopleura and
657 pteropleura of the thorax, white arrows in (A)), which overlaps with the Antp protein
658 pattern ((C') and arrows in (C)). The Antp P2 promoter reporter construct exhibits
659 very weak, if any, expression at these two domains (B-B'). (D-E') Negative controls
660 of Antp protein (D-D') and P1 reporter transcription (E-E') upon overexpression of
661 eGFP. Dashed lines outline the regions of clonal induction in (D) and (E), where
662 neither the Antp protein (D) nor the Antp P1 reporter (E) are repressed. (F-F')
663 Repression of Antp P1 reporter transcription upon clonal overexpression of the full-
664 length untagged Antp protein (Antp-FL). The ectopic expression domain is outlined
665 by white dashed lines in (F) and marked by the expression of eGFP (F'). (G-G')
666 Activation of Antp P1 reporter transcription upon ectopic expression of untagged
667 Antp full-length (Antp-FL) with *Dll* (MD23) driver in the distal region of wing pouch.
668 The ectopic expression domain is outlined by a yellow dashed line in (G) and is
669 marked by the expression of nuclear mRFP1. (H-H') Negative control of ectopic
670 activation of Antp P1 transcription upon overexpression of nuclear mRFP1 alone by
671 *Dll* (MD23)-Gal4. Scale bars denote $100\ \mu m$.
672
673 Figure 2 – figure supplement 2: **Direct correlation between Antp concentration**
674 **and homeotic function – Antp auto-repression and activation occurs at**
675 **endogenous concentrations.** (A-D) Live imaging (one optical section) of
676 SynthAntp-eGFP expressed in the distal antennal portion of the eye-antennal disc by
677 different Gal4 drivers. The concentration was measured using FCS and average

678 concentrations are indicated. An eightfold difference was observed between the
679 strong *Dll*-Gal4 driver (MD23) (A) and weak *69B*-Gal4 driver (D). (E-H)
680 Transformations of the distal antenna into a tarsus in adult flies, caused by
681 *SynthAntp-eGFP* overexpression in antennal discs (A-D). The strength of the
682 transformation correlates with the level of expression from the different Gal4 drivers.
683 Ectopic tarsi range from complete transformation to milder transformations of the
684 arista or ectopic leg bristles in the third antennal segment in (G) and (H), indicated by
685 black arrows. (I) Average FCS measurements performed in nuclei overexpressing
686 *SynthAntp-eGFP*, using different Gal4 drivers. Note that the y-axis amplitudes at the
687 origin of the FCS curves are inversely proportional to the concentration. (J) FCS
688 curves of measurements in (I) normalized to the same amplitude, $G_n(\tau) = 1$ at
689 $\tau = 10\ \mu s$, show major overlap, indicating indistinguishable behavior of Antp binding
690 to chromatin across the concentration range examined $(0.5 - 3.8\ nM)$. (K-P) Antp
691 auto-regulation occurs at endogenous concentrations. (K-L) Repression of
692 endogenous Antp protein upon induction of *SynthAntp-eGFP* in the proximal regions
693 of the wing disc by *69B*-Gal4, which results in Antp expression very similar to
694 endogenous levels. (M-N) No repression is observed upon overexpression of *eGFP*
695 (negative control), as indicated by white arrows in (M). White arrows in (K) and (M)
696 point to the equivalent area in the wing disc, where Antp repression is observed. (O)
697 X-gal stainings of the *Antp* P1 reporter show weak but detectable ectopic β-
698 galactosidase activity in the antennal disc (black arrows). (P) Negative control
699 stainings of eGFP induced by the *69B* enhancer show complete absence of ectopic
700 reporter transcription. Scale bars denote $100\ \mu m$, unless otherwise indicated.
701

702 Figure 2 – figure supplement 3: **Comparison of endogenous and overexpressed**
703 **Antp by FCS.** (A) FCS curves of Antp-eGFP in wing disc nuclei. Concentration
704 differences of fluorescent Antp protein are obvious among cells expressing one or
705 two copies of Antp-eGFP (homozygous and heterozygous larvae) or overexpressing
706 SynthAntp-eGFP from the *Dll* MD23)-Gal4 driver. (B) FCS curves shown in (A)
707 normalized to the same amplitude, $G_n(\tau) = 1$ at $\tau = 10\,\mu s$, show pronounced
708 overlap between homozygous and heterozygous *Antp-eGFP*-expressing cells, as
709 well as between endogenously expressed *Antp* and overexpressed *SynthAntp-*
710 *eGFP*, indicating similar diffusion times and modes of interaction with chromatin.
711 FCS curves are color-coded as outlined in panel (B).
712
713 Figure 3: **Antp is necessary and sufficient to trigger a developmental switch**
714 **from transcriptional auto-activation to auto-repression.** (A and F) Schematic
715 representations of the experimental setup. Clones were induced at 26 h (early
716 clones) or 60 h (late clones). The analysis was performed at third instar larval stage
717 (~96-120 h of development). Black rectangles enclose the blown up regions of
718 interest in the corresponding data panels. (B-B') Early-induced clones, expressing
719 full-length, untagged *Antp*, marked by the absence of mCherry, reveal strong auto-
720 activation of Antp-eGFP (dashed lines in (B)), monitored by the endogenous Antp-
721 eGFP protein. The cyan line in (B) outlines the region of highest endogenous Antp
722 expression. Note the absence of repression of the endogenous gene in
723 overexpression clones within this region at this stage. The whole *Antp* expression
724 domain expresses *Antp-eGFP*, but overexpression clones (sub-regions marked by
725 absence of mCherry staining) express *Antp-eGFP* much stronger (B'). (C-C') Antp is
726 required for sustaining its own expression by auto-activation. Induction of *RNAi*

727 knockdown clones of *Antp* at 26 h of development (early clonal induction, dashed
728 line in (C)) results in cells with pronounced reduction in *Antp* P1 reporter expression.
729 Knockdown clones have been marked by nuclear mRFP1 in (C'). (D) Updated Antp
730 auto-activation model showing the pronounced auto-activating capacity of Antp at
731 early stages, without excluding concurrent weak auto-repression. (E) Study of
732 concentration, DNA-binding and nucleus-to-nucleus variability by FCS at second
733 instar leg and wing discs (derived by the FCS analysis in Figure 3 – figure
734 supplement 3). Low concentration, short decay times (indicative of low degree of
735 DNA-binding) and high variability were observed in second instar wing and leg discs.
736 Reciprocally, high concentration, long decay times (indicative of pronounced
737 chromatin binding) and low variability were observed in third instar wing and leg
738 discs. Statistical significance was determined using a two-tailed Student's T-test (***,
739 $p < 0.001$ and *, $p < 0.05$, namely $p_{concentration}^{2nd\ vs\ 3rd\ instar} = 2.04 \cdot 10^{-20}$, $p_{\tau D2}^{2nd\ vs\ 3rd\ instar} =$
740 $7.2 \cdot 10^{-4}$ and $p_{variation}^{2nd\ vs\ 3rd\ instar} = 3.4 \cdot 10^{-2}$). (G-G') Late-induced clones, expressing
741 full-length, untagged *Antp*, marked by the absence of mCherry, reveal strong auto-
742 repression of Antp-eGFP (dashed lines in (G)), monitored by the endogenous Antp-
743 eGFP protein. The cyan lines in (G) outline the region of strong endogenous
744 expression. (H-H') Late induction of *Antp* knockdown clones by *RNAi* (60 h of
745 development) within its normal expression domain results in upregulation of the *Antp*
746 P1 reporter above endogenous levels (dashed line in (H)). The cyan line in (H)
747 outlines the region of strong endogenous expression of the P1 reporter. Cytoplasmic
748 eGFP marks the clone of *Antp* knockdown by *RNAi* (H'). (I) Updated Antp auto-
749 repression model showing the pronounced auto-repressing capacity of Antp at late
750 stages, without excluding concurrent weak auto-activation.
751

752 Figure 3 – figure supplement 1: **Antp expression patterns are not altered by the**
753 **MiMIC MI02272 insertion.** (A) Schematic representation of the Antp-eGFP fusion
754 protein produced by the conversion of the MiMIC MI02272 construct to an artificial
755 exon. The eGFP-encoding artificial exon is situated in intron 6 of the mRNA and is
756 spliced in between exons 6 and 7 that correspond to the long and non-conserved N-
757 terminal coding sequence of the protein, which has little (if any) function *in vivo*
758 (Papadopoulos et al., 2011), and does not disrupt the homeodomain or YPWM motif.
759 All features have been drawn to scale. (B) Heterozygous flies (embryos and third
760 instar larvae), examined for their Antp-eGFP pattern (detected by an antibody to
761 GFP, green), as compared to the total amount of Antp (expressed by the sum of the
762 MiMIC *Antp-eGFP* and the wild type *Antp* loci), detected by an Antp antibody
763 (magenta). Comparisons of the Antp pattern in wild type embryos and all thoracic
764 imaginal discs are provided case-wise in the right panel. In discs, dashed lines
765 approximately separate the anterior (indicated by "A") from the posterior (indicated
766 by "P") domain of the disc. Note the high expression of Antp in the humeral disc. In
767 the leg discs, Antp is expressed most strongly in the posterior compartment of the
768 prothoracic leg disc, the anterior compartment of the mesothoracic leg disc and in an
769 abundant pattern in the metathoracic leg disc. Cyan arrows point to Antp positive
770 cells in the second and third leg discs that are centrally located, as previously shown
771 (Engstrom et al., 1992). All images represent Z-projections. Scale bars denote
772 $100\ \mu m$.
773
774 Figure 3 – figure supplement 2: **Antp is sufficient and required to trigger a**
775 **developmental switch from transcriptional auto-activation to auto-repression –**
776 **controls.** (A-C') Negative controls of Antp clonal auto-activation and repression

777 using early and late clone induction regimes. (A-A') Without induction of clones
778 expressing full-length untagged Antp, no repression or activation of endogenous
779 Antp protein is observed. (B-C') Upon induction of non-overexpressing clones
780 (clones expressing only Gal4, without a UAS transgene), no activation or repression
781 of Antp protein is observed at early or late induction time points. White dashed lines
782 in (B) and (C) outline the induced clones, marked by the absence of mCherry. (D-E')
783 Early ectopic induction of either Antp full-length untagged protein (D-D') or
784 SynthAntp (E-E') result in upregulation of the Antp P1 reporter. Yellow dashed lines
785 in (D) and arrows in (E) point to the induced clones and cyan continuous lines show
786 the regions of high endogenous expression of the reporter. Clones have been
787 marked by cytoplasmic eGFP. (F-G') Negative controls of early clonal induction of
788 eGFP alone (without concurrent induction of Antp) show no repression of the Antp
789 protein (F-F') or the P1 reporter (G-G'). Dashed lines in (G) mark the clones of eGFP
790 induction. (H-H') Positive control of clonal knockdown of the Antp *RNAi* line used in
791 Figure 3. Clonal knockdown by *RNAi* (indicated by the dashed line in (H) and marked
792 by nuclear mRFP1 in (H')) resulted in efficient downregulation of the endogenous
793 Antp protein. Scale bars denote $100\ \mu m$, unless otherwise indicated.
794
795 Figure 3 – figure supplement 3: **Antp concentration and cell-to-cell variability in**
796 **second and third instar wing and leg imaginal discs** (A-A'') Representative FCS
797 curves recorded in second and third instar wing and leg imaginal discs, expressing
798 *Antp-eGFP*. Note the low concentration in second instar leg and wing discs, reflected
799 by the relatively high amplitude of the FCS curves (inversely proportional to
800 concentration) in (A), as compared to the high concentration in third instar discs in
801 (A'). (B) FCS curves shown in (A) and (A'), normalized to the same amplitude,

802 $G_n(\tau) = 1$ at $\tau = 10\ \mu s$, show a shift towards longer decay times in the third instar
803 leg and wing discs, indicative of pronounced interactions of Antp with chromatin.
804 FCS curves are color-coded as outlined in panel (A). (C) Quantification of average
805 concentrations and cell-to-cell variability in protein concentration among neighboring
806 nuclei in wing and leg, second and third instar, discs. Black bars denote the average
807 concentration and grey bars denote the variability, expressed as the variance over
808 the squared mean. Note the increase in average concentration from second to third
809 instar (eleven-fold increase in the leg disc) and the concurrent drop in variability to
810 almost half of its value. Statistical significance was determined using Student's two-
811 tailed T-test (***, p<0.001, namely $p_{3rd-2nd\ instar\ leg} = 4.4 \cdot 10^{-18}$ and
812 $p_{3rd-2nd\ instar\ wing} = 3.2 \cdot 10^{-8}$).
813
814 Figure 4: **Developmental control of Antp auto-activation and auto-repression**
815 **relies on the relative concentrations of preferentially activating and repressing**
816 **Antp isoforms, which display different binding affinities to chromatin.** (A)
817 Schematic representation of the Antp mRNA, generated from the P1 promoter, with
818 exons represented by grey boxes, the nucleotide length of which is indicated above
819 each exon in base pairs (bp). A higher magnification of exons 4-7 (omitting their in-
820 between splicing points for simplicity) indicates an alternative splice site at the 3' end
821 of exon 7, which results in an isoform featuring a short linker between the YPWM
822 motif and the homeodomain (encoded by four amino acids: RSQF, grey box), or a
823 long linker isoform (encoded by eight amino acids: RSQFGKCQ, white box). All
824 exons have been drawn to scale. (B-B') *SynthAntp-eGFP* transgenes encoding Antp
825 isoforms with a long linker were induced from the *ptc* enhancer and endogenous
826 Antp protein auto-repression was monitored at the proximal portion of the wing disc.

827 A white dashed line outlines the region of auto-repression that was used for
828 quantification (see Materials and Methods). The long linker isoform (preferentially
829 auto-activating isoform) acts as a weak repressor of Antp protein, as compared to its
830 short linker equivalent in (F-F'). (C-C') *SynthAntp-eGFP* transgenes encoding Antp
831 isoforms with a long linker were induced from the *Dll* (MD23) enhancer and Antp P1
832 reporter auto-activation was monitored at the distal portion of the wing disc (wing
833 blade), indicated by a yellow dashed line. The long linker isoform (preferential auto-
834 activator) induces stronger transcription of the P1 Antp reporter, as compared to its
835 short linker counterpart in (G-G'). (D) Relative concentration of long linker isoform
836 (see Materials and Methods); quantification of its auto-activation and auto-repression
837 capacities (see Materials and Methods); fraction of DNA-bound fluorescent
838 molecules, measured by FCS (see also Figure 4 – figure supplement 2); and relative
839 affinity of binding to the DNA (quantified by the calculation of its apparent equilibrium
840 dissociation constant using both FCS measurements and gel-shifts (EMSA), see
841 also Figure 4 – figure supplements 2 and 3) are provided for comparison with (H).
842 (E) Updated qualitative model of the dual auto-regulation of Antp by activating and
843 repressing Antp isoforms, equivalent to an early developmental stage. The activating
844 isoform binds with lower affinity to the P1 Antp promoter, but is produced in excess,
845 relative to the repressing one, resulting in preferential activation of transcription,
846 without cancelling out partial binding of the auto-repressive isoform as well. (F-F')
847 Similar to (B-B'), except induction has been with the short linker (preferentially auto-
848 repressing) isoform. It acts as a stronger repressor of Antp protein, as compared to
849 its long linker equivalent in (B-B'). (G-G') Similar to (C-C'), except induction has been
850 with the short linker (preferentially auto-repressing) isoform. It acts as a weaker
851 activator of Antp P1 transcription, as compared to its long linker equivalent in (C-C').

852 (H) The same parameters as in (D) have been evaluated and are provided for
853 comparison with (D). Statistical significance was determined using a two-tailed
854 Student's T-test between measurements performed with the long linker (auto-
855 activating) isoform in (D) and the short linker (auto-repressing) isoform in (H) (***,
856 $p < 0.001$ and *, $p < 0.05$, namely $p_{concentration}^{embryo\ long\ vs\ short} = 3.16 \cdot 10^{-5}$,
857 $p_{concentration}^{2nd\ instar\ long\ vs\ short} = 1.16 \cdot 10^{-4}$, $p_{concentration}^{3rd\ instar\ long\ vs\ short} = 2.85 \cdot 10^{-6}$,
858 $p_{relative\ activation}^{long\ vs\ short} = 4.1 \cdot 10^{-3}$, $p_{relative\ repression}^{long\ vs\ short} = 2.4 \cdot 10^{-4}$ and
859 $p_{DNA-bound\ fraction\ (FCS)}^{long\ vs\ short} = 5.6 \cdot 10^{-10}$). (I) Updated qualitative model representation of
860 Antp repression as in (E), whereby at later stages excess of Antp auto-repressor
861 accounts for negative feedback on transcriptional regulation of the P1 promoter,
862 without ruling out partial binding by the activating isoform as well, resulting in partial
863 activation of transcription, hence expression is maintained.
864
865 Figure 4 – supplement 1: **Developmental control of Antp auto-activation and**
866 **repression relies on the relative concentrations of preferentially auto-**
867 **activating and auto-repressing Antp isoforms, which display different binding**
868 **affinities to chromatin – short linker isoform controls.** (A-H') Experiments of
869 Figures 2-4, performed with short linker (preferentially auto-repressing) full-length
870 and SynthAntp isoforms on their capacity to repress and activate Antp P1 reporter
871 transcription and Antp protein. Dashed lines in all panels outline the clones induced
872 or the region of ectopic expression using *Dll* (MD23)-Gal4, whereas closed
873 continuous cyan lines outline the regions of endogenous Antp P1 reporter
874 expression in (G) and (H). (A-A') Repression of Antp protein by late clonal induction
875 of SynthAntp in the wing notum. (B-B') Equivalent assay as in (A-A'), but monitoring
876 auto-repression of the Antp P1 promoter transcription. (C-C') Similar assay to (B-B'),

877 using the full-length Antp protein, induced at the later time point. (D-D') Ectopic
878 induction of full-length, short linker, untagged Antp cDNA with concurrent labeling of
879 the expression domain by nuclear mRFP1 results in weak ectopic auto-activation of
880 the Antp P1 reporter. (E-F') Early and late clonal induction of full-length, short linker,
881 untagged Antp results in auto-repression (E-E'), or induction (F-F'), of the
882 endogenous Antp protein, respectively. (G-H') Early clonal induction of SynthAntp
883 (G-G') or the full-length cDNA (H-H'), both featuring a short linker, triggers ectopic
884 activation of P1 promoter transcription. (I-J') Antp long (I-I') and short (J-J') linker
885 isoforms repress Antp at the transcriptional level (monitored by Antp P1 reporter
886 expression) when induced by *ptc*-Gal4 in the wing disc. Arrows point to the regions
887 of auto-repressed Antp promoter. Scale bars denote $100\,\mu m$, unless otherwise
888 indicated.
889
890 Figure 4 – figure supplement 2: **Comparative binding study of Antp short and**
891 **long linker isoforms by FCS.** (A-B) FCS analysis performed on third instar wing
892 and antennal imaginal discs, expressing short or long linker Antp isoforms (tagged to
893 eGFP) close to endogenous concentrations, from the *69B*-enhancer. Cell nuclei of
894 similar concentrations in the two datasets have been selected for analysis (A).
895 Average FCS measurements on the short linker Antp isoform display a consistent
896 shift towards longer decay times, as compared to its long linker counterpart (B),
897 indicating higher degree of chromatin binding. (C-D') Binding study of short and long
898 linker Antp isoforms in third instar wing and antennal discs, expressed by *69B*-Gal4.
899 The concentration of the Antp short and long linker isoform DNA-bound complexes
900 (derived by fitting the FCS curves in (A)) is plotted as a function of the total
901 concentration of Antp-eGFP molecules. From the linear regression equations,

902 $y = 0.34x - 5.31$ (D') and $y = 0.24x - 3.28$ (E'), the ratio of apparent dissociation
903 constants for the long and short linker isoforms was calculated to be $\frac{K_{d,Antp}^{long\ linker\ isof.}}{K_{d,Antp}^{short\ linker\ isof.}} >$
904 2.3 (for the calculation refer to Supplement 3). The two dissociation constants differ
905 at least 2.3 times, indicating stronger binding of the short linker isoform to the DNA,
906 as compared to the long linker one.
907
908 Figure 4 – figure supplement 3: **In vitro binding study of Antp full-length long**
909 **and short linker isoforms to Antp and homeodomain binding sites by gel-shift**
910 **assays.** (A-D) Gel-shift experiments using purified full-length Antp protein, featuring
911 a long or a short linker, show from twofold (C) to eightfold (A) stronger binding of the
912 short linker isoform to Antp binding sites. The Antp-DNA complexes and the free
913 DNA-binding site are shown in (A) and are the same for (B-D). In (C) and (D) an
914 antibody to Antp has been used in the fifth lane of the serial dilutions of Antp protein
915 (black asterisks) to indicate specific Antp binding (super-shift). Equal starting
916 concentrations of purified Antp short and long linker protein have been used and
917 determined by Western blot. Each successive lane of long or short linker isoform
918 binding represents a twofold lower concentration of Antp protein. Triangles above the
919 concentration range of each sample denote declining concentrations and the black
920 areas therein indicate the corresponding concentration range at which Antp binding
921 to the probe has been observed.
922
923 Figure 5: **Mathematical model for stochastic Antp expression and system**
924 **perturbations.** (A-A'') A dynamic promoter, which drives transcription of *Antp*
925 followed by a splicing step, leads to either the repressing ("R" in (A)) or activating
926 ("A" in (A)) isoform of Antp. In the absence of Antp, the promoter is inactive and

927 transcription cannot take place (“U” in (A)). This promoter configuration leads to
928 suppression of variability (A’) and increase in Antp concentration (A’’). (B-B’’)
929 Competition of Antp binding, whereby state “A” can be reached only through the
930 unbound state “U” in (B), results in increase in Antp protein numbers (B’’) without
931 decrease in variability (cyan bar in (B’)). (C-C’’) Requirement of the negative
932 feedback for suppression of variability. In the absence of the state “R” (C),
933 concentration increases (C’’), but variability also increases rather than being
934 suppressed (cyan bar in (C’)). (D-D’) In all three models (A-A’’, B-B’’ and C-C’’),
935 changing relative concentrations of repressing and activating Antp isoforms are
936 predicted, starting with excess of the activating isoform at early stages (D) and
937 ending with excess of the repressing one at later stages (D’). (E-F’) Model
938 predictions (E and F) and experimental data validation (E’ and F’) of the decrease of
939 variability (E) and the increase of the protein Fano factor (F) as a function of
940 increasing Antp concentrations. (G-K) Perturbations of the model in (A-A’’) by
941 overexpression of Antp long linker (auto-activating) and short linker (auto-repressing)
942 isoforms. (G-H’) High concentration of both auto-repressing (G-G’) and auto-
943 activating (H-H’) Antp isoforms result in increase of Antp protein copy numbers in
944 (G’) and (H’) with concurrent repression of variability in (G) and (H), similar to the
945 unperturbed system in (A-A’’). (I-K) Overexpression of SynthAntp activating or
946 repressing isoforms by *Dll* (MD23)-Gal4 in distal portion of appendages, such as the
947 antenna and the leg imaginal discs, result in tarsal transformations in the antenna
948 (where Antp is normally not expressed (I)), but normal development of tarsi in all legs
949 (J). These flies are fully viable and can be maintained as a stock. Similarly normal is
950 the development of the prescutum region of the notum, upon overexpression of
951 SynthAntp activating or repressing isoforms by *MS243*-Gal4 (K). (L-O) Perturbations

952 of the model system in (A-A'') by overexpression of an exogenous Antp repressor.
953 As expected, overexpression of the Hox transcription factor Scr, which is able to only
954 repress, but neither activate Antp expression, nor its own (see Figure 5 – figure
955 supplement 1 B-E' for details), is predicted to result in drastic reduction of Antp at
956 later stages (cyan bar in (L')) and have barely any effect in Antp variability at these
957 stages (L). In this case, ectopic expression of SynthScr by *Dll* (MD23)-Gal4 still
958 triggers transformation of the antenna, which is devoid of Antp expression, into a
959 tarsus (M), similar to induction of Antp in this tissue (I), but in overlapping domains
960 with Antp normal expression, such as the distal leg portion, malformations of the
961 tarsus and femur are observed (N). Overexpression of SynthScr transgenes by
962 *MS243*-Gal4 in the fly notum results in severe malformations, indicated by
963 developmental defects of the adult cuticle (O). Both fly genotypes in (N) and (O) die
964 as pharate adults.
965
966 Figure 5 – figure supplement 1: **Investigation of extrinsic variability in the**
967 **endogenous Antp-eGFP expression domain.** (A-B') Live imaging of a wing disc
968 notum, where nuclear mRFP1 protein is highly expressed from a constitute enhancer
969 (*ubi*-mRFP1(NLS)), alongside with endogenous Antp-eGFP. FCS measurements
970 were performed in the region of high co-expression of Antp and RFP. (B-B') Higher
971 magnification of cells as in (A-A'). Note the uneven distribution of Antp in the nuclei
972 and the formation of sites of accumulation in (B). (C) Plot of the concentration of
973 Antp-eGFP (expressed as number of fluorescent molecules in the Observation
974 Volume Element (OVE)). The correlation coefficient r was calculated to be $r = 0.523$
975 and the p-value to be $p_{correlation} = 9.77 \cdot 10^{-5}$. Scale bars denote $100\,\mu m$, unless
976 otherwise indicated.

977

978 Figure 5 – figure supplement 2: **Controls of Antp model predictions and Scr-**
979 **mediated perturbations.** (A-A') *MS243*-Gal4-mediated expression of repressing or
980 activating SynthAntp isoforms results in repression of the Antp endogenous protein
981 in the notum region of the wing disc. (B-B') Ectopic expression of SynthScr in the
982 wing disc using *ptc*-Gal4 results in drastic reduction of endogenous Antp protein
983 levels. (C-C') Ectopic expression of SynthScr by *MS243*-Gal4 in the notum results in
984 repression of the Antp protein. (D-D') SynthScr represses Antp at the transcriptional
985 level, as indicated by the absence of transcription of the Antp P1 reporter (white
986 arrows in (D)). (E-E') Unlike SynthAntp, SynthScr is not able to activate the Antp P1
987 promoter reporter transcription (E), when induced by *Dll* (MD23)-Gal4. (F-F')
988 SynthScr is not able to downregulate its own endogenous protein levels upon
989 overexpression by *Dll* (MD23)-Gal4. Dashed line in (F') outlines the region of high
990 overlap between the overexpressed SynthScr and endogenous Scr stainings. (G-G')
991 Negative control staining for the induction of eGFP in the wing disc notum by
992 *MS243*-Gal4, which fails to repress endogenous Antp protein. Dashed lines in (A),
993 (B) and (C) outline the regions of ectopic overexpression of SynthAntp or SynthScr,
994 where endogenous Antp is repressed, whereas the dashed line in (F') outlines the
995 region of overlap between SynthScr overexpression and endogenous expression of
996 the Scr protein, where no repression is observed. Scale bars denote $100\ \mu m$, unless
997 otherwise indicated.

998

999 ## Acknowledgements

1000

1001 We are deeply saddened by the unexpected passing of Prof. Walter J.
1002 Gehring, at the very inception of this work, when the project was still in the planning
1003 and preliminary data gathering stage. Prof. Gehring was an extraordinary human
1004 being and a scientific giant, whose work will continue to educate and inspire
1005 generations to come. The authors are indepted to Sonal Nagarkar Jaiswal, Paolo
1006 Mangahas, and Hugo J. Bellen for creating and sharing with us the Antp-eGFP line.
1007 DKP has been supported by a long-term fellowship from the Swiss National Science
1008 Foundation (PBBSP-138700) and a long-term fellowship from the Federation of
1009 European Biochemical Societies (FEBS) at initial stages of this project. VV has been
1010 supported by the Knut and Alice Wallenberg foundation and Karolinska Institute
1011 Research Funds. DKP would like to express his gratitude to PT for outstanding
1012 scientific, and uninterrupted financial, support. DKP is indebted to Markus Burkhardt,
1013 head of the imaging platform at the Center for Regenerative Therapies Dresden
1014 (CRTD), for help and discussions regarding FCS experiments in Dresden; Sylke
1015 Winkler and the DNA sequencing facility of the Max-Planck Institute of Molecular Cell
1016 Biology and Genetics (MPI-CBG) for providing assistance with Antp transcript variant
1017 quantification; Aliona Bogdanova and the Protein expression and purification facility
1018 of the MPI-CBG for purification of the repressing and activating Antp isoforms; as
1019 well as the Light Microscopy facility of MPI-CBG. DKP is also grateful to KS for
1020 numerous discussions and support throughout the implementation of this work and
1021 to Konstantinos Papadopoulos for advice on the mathematical analysis of the
1022 relative binding constants of repressing and activating Antp isoforms. The authors
1023 would like to acknowledge Jan Brugues and Thomas M. Schultheiss for their critical
1024 reading and insightful comments on the manuscript.
1025

1026 ## Materials and Methods

1027

1028 Fly stocks used

1029 The Antp-eGFP MiMIC line has been a kind gift from Hugo J. Bellen. The
1030 *atonal* (VDRC ID 318959), *brinker* (VDRC ID 318246), *spalt major* (VDRC ID
1031 318068), *yorkie* (VDRC ID 318237), *senseless* (VDRC ID 318017) and *Sex combs*
1032 *reduced* (VDRC ID 318441) fosmid lines are available from the Vienna Drosophila
1033 Resource Center (VDRC) and have been generated recently in our laboratory (Sarov
1034 et al., 2016). The *fork head* (stock 43951), *grainy head* (stock 42272), *Abdominal B*
1035 (stock 38625), *eyeless,* (stock 42271), *spineless* (transcript variant A, stock 42289),
1036 and *grain* (stock 58483) tagged BACs were generated by Rebecca Spokony and
1037 Kevin P. White and are available at the Bloomington Stock Center. For the *scalloped*
1038 gene, a GFP-trap line was used (Buszczak et al., 2007), a kind gift from Allan C.
1039 Spradling laboratory (line CA07575), with which genome-wide chromatin
1040 immunoprecipitation experiments have been performed (Slattery et al., 2013). For
1041 the *spineless* gene, Bloomington stock 42676, which tags isoforms C and D of the
1042 Spineless protein has been also tried in fluorescence imaging and FCS experiments,
1043 but did not yield detectable fluorescence in the antennal disc, rendering it
1044 inappropriate to be used in our analysis. Therefore, we resided to stock 42289,
1045 which tags the A isoform of the protein. For the *eyeless* gene, the
1046 FlyFos015860(pRedFlp-Hgr)(ey13630::2XTY1-SGFP-V5-preTEV-BLRP-
1047 3XFLAG)dFRT line (VDRC ID 318018) has been tried also in fluorescence imaging
1048 and FCS experiments, but did not yield detectable fluorescence in the eye disc for it
1049 to be used in our analysis. The *act5C*-FRT-yellow-FRT-Gal4 (Ay-Gal4) line used for
1050 clonal overexpression or *RNAi* knockdown has been described (Ito et al., 1997). The

1051 UAS-Antp lines (synthetic and full-length), as well as UAS-SynthScr constructs have
1052 been previously described (Papadopoulos et al., 2011; Papadopoulos et al., 2010).
1053 The *Dll*-Gal4 (MD23) line has been a kind gift of Ginés Morata (Calleja et al., 1996).
1054 *69B*-Gal4 and *ptc*-Gal4 have been obtained from the Bloomington Stock Center. The
1055 *Antp* P1-*lacZ* and P2-*lacZ* have been previously described (Engstrom et al., 1992;
1056 Zink et al., 1991). The P1 reporter construct spans the region between 9.4 kb
1057 upstream of the P1 promoter transcription initiation site and 7.8 kb downstream into
1058 the first intron, including the first exon sequences and thus comprising 17.2 kb of
1059 Antp regulatory sequences (pAPT 1.8). The line used has been an insertion of the
1060 pAPT 1.8 vector bearing the P1 promoter regulatory sequences upstream of an
1061 *actin-lacZ* cytoplasmic reporter and has been inserted in cytogenetic location 99F on
1062 the right chromosomal arm of chromosome 3. The *Antp-RNAi* line has been from
1063 VDRC, line KK101774. UAS-eGFP stock was a kind gift of Konrad Basler. We are
1064 indebted to Sebastian Dunst for generating the *ubi*-FRT-mCherry(stop)-FRT-
1065 Gal4(VK37)/CyO line, which drives clonal overexpression upon flippase excision,
1066 while simultaneously marking cells by the loss of mCherry. For red-color labeling of
1067 clones the *act5C*-FRT-CD2-FRT-Gal4, UAS-mRFP1(NLS)/TM3 stock 30558 from
1068 the Bloomington Stock Center has been used. For marking the ectopic expression
1069 domain of untagged Antp proteins the UAS-mRFP1(NLS)/TM3 stock 31417 from the
1070 Bloomington Stock Center has been used. The *MS243*-Gal4; UAS-GFP/CyO line
1071 was a kind gift from the laboratory of Ernesto Sánchez-Herrero.
1072
1073 <u>Fly genotypes corresponding to fluorescence images</u>
1074 Figure 1 – figure supplement 1 A: FlyFos018487(pRedFlp-Hgr)(ato37785::2XTY1-
1075 SGFP-V5-preTEV-BLRP-3XFLAG)dFRT

1076 Figure 1 – figure supplement 1 B: FlyFos024884(pRedFlp-Hgr)(brk25146::2XTY1-
1077 SGFP-V5-preTEV-BLRP-3XFLAG)dFRT
1078 Figure 1 – figure supplement 1 C: FlyFos030836(pRedFlp-Hgr)(salm30926::2XTY1-
1079 SGFP-V5-preTEV-BLRP-3XFLAG)dFRT
1080 Figure 1 – figure supplement 1 D: FlyFos029681(pRedFlp-Hgr)(yki19975::2XTY1-
1081 SGFP-V5-preTEV-BLRP-3XFLAG)dFRT
1082 Figure 1 – figure supplement 1 E: w^{1118}; PBac(fkh-GFP.FPTB)VK00037/SM5
1083 Figure 1 – figure supplement 1 F: *sd*-eGFP (FlyTrap, homozygous)
1084 Figure 1 – figure supplement 1 G: w^{1118}; PBac(grh-GFP.FPTB)VK00033
1085 Figure 1 – figure supplement 1 H: FlyFos018974(pRedFlp-Hgr)(Scr19370::2XTY1-
1086 SGFP-V5-preTEV-BLRP-3XFLAG)dFRT
1087 Figure 1 – figure supplement 1 I: FlyFos015942(pRedFlp-Hgr)(sens31022::2XTY1-
1088 SGFP-V5-preTEV-BLRP-3XFLAG)dFRT
1089 Figure 1 – figure supplement 1 J and K: Antp-eGFP (MiMIC) homozygous (line
1090 MI02272, converted to an artificial exon)
1091 Figure 1 – figure supplement 1 L: w^{1118}; PBac(Abd-B-EGFP.S)VK00037/SM5
1092 Figure 1 – figure supplement 1 M: w^{1118}; PBac(ey-GFP.FPTB)VK00033
1093 Figure 1 – figure supplement 1 N: w^{1118}; PBac(ss-GFP.A.FPTB)VK00037
1094 Figure 1 – figure supplement 1 O and P: w^{1118}; PBac(grn-GFP.FPTB)VK00037
1095 Figure 2 B and B': *hs*-flp/+; *act5C*-FRT-yellow-FRT-Gal4/+; UAS-SynthAntp long
1096 linker-eGFP/+
1097 Figure 2 C and C': *hs*-flp/+; *act5C*-FRT-yellow-FRT-Gal4, UAS-eGFP/+; UAS-Antp
1098 long linker (full-length, untagged)/+
1099 Figure 2 F and F': *Dll*-Gal4 (MD23)/+; UAS-SynthAntp-eGFP/*Antp* P1-*lacZ*
1100 Figure 2 – figure supplement 1 A and A': *Antp* P1-*lacZ*/TM3

1101 Figure 2 – figure supplement 1 B and B': *Antp* P2-*lacZ*/CyO

1102 Figure 2 – figure supplement 1 C and C': wild type

1103 Figure 2 – figure supplement 1 D and D': *hs*-flp; *act5C*-FRT-yellow-FRT-Gal4, UAS-

1104 eGFP

1105 Figure 2 – figure supplement 1 E and E': *hs*-flp/+; *act5C*-FRT-yellow-FRT-Gal4,

1106 UAS-eGFP/+; *Antp* P1-*lacZ*/+

1107 Figure 2 – figure supplement 1 F and F': *hs*-flp/+; *act5C*-FRT-yellow-FRT-Gal4, UAS-

1108 eGFP/+; UAS-Antp long linker (full-length, untagged)/*Antp* P1-*lacZ*

1109 Figure 2 – figure supplement 1 G and G': *Dll*-Gal4 (MD23)/+; UAS-Antp long linker

1110 (full-length, untagged), UAS-mRFP1(NLS)/ *Antp* P1-*lacZ*

1111 Figure 2 – figure supplement 1 H and H': *Dll*-Gal4 (MD23)/+; UAS-mRFP1(NLS)/

1112 *Antp* P1-*lacZ*

1113 Figure 2 – figure supplement 2 A and E: *Dll*-Gal4 (MD23)/+; UAS-SynthAntp long

1114 linker-eGFP/+

1115 Fig 2 – figure supplement 2 B and F: *ptc*-Gal4/+; UAS-SynthAntp long linker-eGFP/+

1116 Figure 2 – figure supplement 2 C and G: *Dll*-Gal4 (MD713)/+; UAS-SynthAntp long

1117 linker-eGFP/+

1118 Figure 2 – figure supplement 2 D, H and K, L, O: *69B*-Gal4/UAS-SynthAntp long

1119 linker-eGFP

1120 Figure 2 – figure supplement 2 M, N and P: *69B*-Gal4/UAS- eGFP

1121 Figure 3 B, B', G and G': *hs*-flp/+; *ubi*-FRT-mChery-FRT-Gal4/+; Antp-eGFP

1122 (MiMIC)/UAS-Antp long linker (full-length, untagged)

1123 Figure 3 C and C': *hs*-flp/+; UAS-Antp^{RNAi}/+; *Antp* P1-*lacZ*/*act5C*-FRT-CD2-FRT-

1124 Gal4, UAS-mRFP1(NLS)

1125 Figure 3 H and H': *hs*-flp/+; UAS-Antp^{RNAi}/*act5C*-FRT-yellow-FRT-Gal4, UAS-eGFP;
1126 *Antp* P1-*lacZ*/+
1127 Figure 3 – figure supplement 1 B: *Antp* P1-*lacZ*/TM6B
1128 Figure 3 – figure supplement 2 A and A': *hs*-flp/+; *ubi*-FRT-mChery-FRT-Gal4/+;
1129 Antp-eGFP (MiMIC)/UAS-Antp long linker (full-length, untagged)
1130 Figure 3 – figure supplement 2 B-C': *hs*-flp/+; *ubi*-FRT-mChery-FRT-Gal4/+; Antp-
1131 eGFP (MiMIC)/+
1132 Figure 3 – figure supplement 2 D and D': *hs*-flp/+; *act5C*-FRT-yellow-FRT-Gal4,
1133 UAS-eGFP/+; *Antp* P1-*lacZ*/UAS-Antp long linker (full-length, untagged)
1134 Figure 3 – figure supplement 2 E and E': *hs*-flp/+; *act5C*-FRT-yellow-FRT-Gal4/+;
1135 UAS-SynthAntp long linker-eGFP/+
1136 Figure 3 – figure supplement 2 F and F': *hs*-flp/+; *act5C*-FRT-yellow-FRT-Gal4, UAS-
1137 eGFP/+
1138 Figure 3 – figure supplement 2 G and G': *hs*-flp/+; *act5C*-FRT-yellow-FRT-Gal4,
1139 UAS-eGFP/+; *Antp* P1-*lacZ*/+
1140 Figure 3 – figure supplement 2 H and H': *hs*-flp/+; UAS-Antp^{RNAi}/+; Antp-eGFP
1141 (MiMIC)/*act5C*-FRT-CD2-FRT-Gal4, UAS-mRFP1(NLS)
1142 Figure 4 B and B': *ptc*-Gal4/+; UAS-SynthAntp long linker-eGFP/+
1143 Figure 4 C and C': *Dll*-Gal4 (MD23)/+; UAS-SynthAntp long linker-eGFP/*Antp* P1-
1144 *lacZ*
1145 Figure 4 F and F': *ptc*-Gal4/+; UAS-SynthAntp long linker-eGFP/+
1146 Figure 4 G and G': *Dll*-Gal4 (MD23)/+; UAS-SynthAntp short linker-eGFP/*Antp* P1-
1147 *lacZ*
1148 Figure 4 – figure supplement 1 A and A': *hs*-flp/+; *act5C*-FRT-yellow-FRT-Gal4/+;
1149 UAS-SynthAntp short linker-eGFP/+

1150 Figure 4 – figure supplement 1 B, B', G and G': *hs*-flp/+; *act5C*-FRT-yellow-FRT-
1151 Gal4/+; UAS-SynthAntp short linker-eGFP/*Antp* P1-*lacZ*
1152 Figure 4 – figure supplement 1 C, C', H and H': *hs*-flp/+; *act5C*-FRT-yellow-FRT-
1153 Gal4/+; UAS-Antp short linker (full-length, untagged)/*Antp* P1-*lacZ*
1154 Figure 4 – figure supplement 1 D and D': *hs*-flp/+; *Dll*-Gal4 (MD23)/+; UAS-Antp
1155 short linker (full-length, untagged), UAS-mRFP1(NLS)/*Antp* P1-*lacZ*
1156 Figure 4 – figure supplement 1 E-F': *hs*-flp/+; *ubi*-FRT-mChery-FRT-Gal4/+; Antp-
1157 eGFP (MiMIC)/UAS-Antp short linker (full-length, untagged)
1158 Figure 4 – figure supplement 1 I and I': *ptc*-Gal4/+; UAS-SynthAntp long linker-
1159 eGFP/*Antp* P1-*lacZ*
1160 Figure 4 – figure supplement 1 J and J': *ptc*-Gal4/+; UAS-SynthAntp short linker-
1161 eGFP/*Antp* P1-*lacZ*
1162 Figure 5 I and J: *Dll*-Gal4 (MD23)/+; UAS-SynthAntp long linker-eGFP/+ or *Dll*-Gal4
1163 (MD23)/+; UAS-SynthAntp short linker-eGFP/+
1164 Figure 5 K: *MS243*-Gal4/+; UAS-SynthAntp long linker-eGFP/*Dr* or *MS243*-Gal4/+;
1165 UAS-SynthAntp long linker-eGFP/*Dr*
1166 Figure 5 M and N: *Dll*-Gal4 (MD23)/+; UAS-mCitrine-SynthScr/+
1167 Figure 5 O: *MS243*-Gal4/+; UAS-mCitrine-SynthScr/+
1168 Figure 5 P: *MS243*-Gal4/+; UAS-mCitrine-SynthScr/+
1169 Figure 5 – figure supplement 1 A-B': *ubi*-mRFP1(NLS)/+ or y; Antp-eGFP (MiMIC)/+
1170 Figure 5 – figure supplement 2 A and A': *MS243*-Gal4/+; UAS-SynthAntp long linker-
1171 eGFP/*Dr* or *MS243*-Gal4/+; UAS-SynthAntp long linker-eGFP/*Dr*
1172 Figure 5 – figure supplement 2 B and B': *ptc*-Gal4/+; UAS-mCitrine-SynthScr/+
1173 Figure 5 – figure supplement 2 C and C': *MS243*-Gal4/+; UAS-mCitrine-SynthScr/+

1174 Figure 5 – figure supplement 2 D and D': *ptc*-Gal4/+; UAS-mCitrine-SynthScr/*Antp*
1175 P1-*lacZ*
1176 Figure 5 – figure supplement 2 E and E': *Dll*-Gal4 (MD23)/+; UAS-mCitrine-
1177 SynthScr/*Antp* P1-*lacZ*
1178 Figure 5 – figure supplement 2 F and F': *Dll*-Gal4 (MD23)/+; UAS-mCitrine-
1179 SynthScr/+
1180 Figure 5 – figure supplement 2 G and G': *MS243*-Gal4/+; UAS-eGFP/+
1181
1182 Preparation of second and third instar imaginal discs for FCS measurements
1183 For FCS measurements, imaginal discs (eye-antennal, wing, leg, humeral and
1184 genital) and salivary glands were dissected from third instar wandering larvae, or
1185 wing and leg discs from second instar larvae, in Grace's insect tissue culture
1186 medium (ThermoFisher Scientific, 11595030) and transferred to 8-well chambered
1187 coverglass (Nunc® Lab-Tek™, 155411) containing PBS just prior to imaging or FCS
1188 measurements. Floating imaginal discs or salivary glands were sunk to the bottom of
1189 the well using forceps.
1190
1191 Immunostainings in larval imaginal discs
1192 Larval imaginal discs were stained according to (Papadopoulos et al., 2010).
1193 Stainings for the endogenous Antp protein have been performed using a mouse anti-
1194 Antp antibody (Developmental Studies Hybridoma Bank, University of Iowa, anti-
1195 Antp 4C3) in a dilution of 1:250 for embryos and 1:500 for imaginal discs. eGFP, or
1196 eGFP-tagged proteins have been stained using mouse or rabbit anti-GFP antibodies
1197 from ThermoFisher Scientific in a dilution of 1:500 in imaginal discs and 1:250 in
1198 embryos. mRFP1 was stained using a Chromotek rat anti-RFP antibody. For *Antp*

1199 P1 promoter stainings in imaginal discs we used the mouse anti-β-galactosidase 40-
1200 1a antibody from Developmental Studies Hybridoma Bank, University of Iowa in a
1201 dilution of 1:50. The rabbit anti-Scr antibody was used in a dilution of 1:300 (LeMotte
1202 et al., 1989). Confocal images of antibody stainings represent predominatly Z-
1203 projections and Zeiss LSM510, Zeiss LSM700 or Zeiss LSM880 Airyscan confocal
1204 laser scanning microscopy systems with an inverted stand Axio Observer
1205 microscope were used for imaging. Image processing and quantifications have been
1206 performed in Fiji (Schindelin et al., 2012). For optimal spectral separation, secondary
1207 antibodies coupled to Alexa405, Alexa488, Alexa594 and Cy5 (ThermoFischer
1208 Scientific) were used.
1209
1210 Colocalization of wild type and eGFP-tagged MiMIC Antp alleles in imaginal discs
1211 To examine whether the pattern of the MiMIC Antp-eGFP fusion protein
1212 recapitulates the Antp wild type expression pattern in both embryo and larval
1213 imaginal discs, we performed immunostainings of heterozygous Antp-eGFP and wild
1214 type flies to visualize the embryonic (stage 13) and larval expression of *Antp* and
1215 eGFP. In this experiment, we 1) visualized the overlap between eGFP and *Antp* (the
1216 eGFP pattern reflects the protein encoded by the MiMIC allele, whereas the *Antp*
1217 pattern reflects the sum of protein produced by the MiMIC allele and the allele of the
1218 balancer chromosome) and 2) compared the eGFP expression pattern to the Antp
1219 expression pattern in wild type discs and embryos.
1220
1221 Induction of early and late overexpression and *RNAi*-knockdown clones in imaginal
1222 discs

1223 Genetic crosses with approximately 100 virgin female and 100 male flies were
1224 set up in bottles and the flies were allowed to mate for 2 days. Then, they were
1225 transferred to new bottles and embryos were collected for 6 hours at 25°C. Flies
1226 were then transferred to fresh bottles and kept until the next collection at 18°C. To
1227 asses Antp auto-activation, the collected eggs were allowed to grow at 25°C for 26 h
1228 from the midpoint of collection, when they were subjected to heat-shock by
1229 submersion to a water-bath of 38°C for 30 min and then placed back at 25°C until
1230 they reached the stage of third instar wandering larvae, when they were collected for
1231 dissection, fixation and staining with antibodies. To assess Antp auto-repression, the
1232 same procedure was followed, except that the heat-shock was performed at 60 h of
1233 development after the midpoint of embryo collection. Whenever necessary, larval
1234 genotypes were selected under a dissection stereomicroscope with green and red
1235 fluorescence filters on the basis of *deformed* (*dfd*)-YFP bearing balancer
1236 chromosomes (Le et al., 2006) and visual inspection of fluorescence in imaginal
1237 discs.
1238
1239 Measurement of Antp transcript variant abundance
1240 The linker between the Antp YPWM motif and the homeodomain contains the
1241 sequence RSQFGKCQE. Short linker isoforms encode the sequence RSQFE,
1242 whereas long linker isoforms are generated by alternative splicing of a 12 base pair
1243 sequence encoding the four amino acid sequence GKCQ into the mRNA. We initially
1244 designed primer pairs for RT-qPCR experiments to distinguish between the short
1245 and long linker mRNA variants. For the short linker variant, we used nucleotide
1246 sequences corresponding to RSQFERKR (with RKR being the first 3 amino acids of
1247 the homeodomain). For detection of the long linker variant we designed primers

1248 either corresponding to the RSQFGKCQ sequence, or GKCQERKR. We observed in
1249 control PCRs (using plasmid DNA harboring either a long or a short linker cDNA)
1250 that primers designed for the short linker variant still amplified the long linker one.
1251 Moreover, with linker sequences differing in only four amino acids, encoded by 12
1252 base pars, primer pairs flanking the linker could also not be used, since, due to very
1253 similar sizes, both variants would be amplified in RT-qPCR experiments with almost
1254 equal efficiencies. Therefore, we used primer pairs flanking the linker region to
1255 indiscriminately amplify short and long linker variants, using non-saturating PCR (18
1256 cycles) on total cDNA generated from total RNA. We then resolved and assessed
1257 the relative amounts of long and short linker amplicons in a second step using
1258 Fragment Analyzer (Advanced Analytical). RNA was extracted from stage 13
1259 embryos, second instar larvae at 60 h of development, and leg or wing discs from
1260 third instar wandering larvae using the Trizol® reagent (ThermoFischer Scientific),
1261 following the manufacturer's instructions. Total RNA amounts were measured by
1262 NanoDrop and equal amounts were used to synthesize cDNA using High-Capacity
1263 RNA-to-cDNA™ Kit (ThermoFischer Scientific), following the manufacturer's
1264 instructions. Total cDNA yields were measured by NanoDrop and equal amounts
1265 were used in PCR, using in-house produced Taq polymerase. 10 ng of plasmid DNA,
1266 bearing either a long or a short transcript cDNA were used as a control. PCR product
1267 abundance was analyzed both by agarose gel electrophoresis and using Fragment
1268 Analyzer (Advanced Analytical).

1269 The quantification of the transcript variant concentration (Figure 4 D and H)
1270 has been made considering 100% (value equal to 1 on the y axis) as the sum of long
1271 and short isoforms at each developmental stage, whereas the quantification of the
1272 relative activation and repression efficiency has been performed considering the

1273 short linker variant as having 100% repression and the long linker variant as having
1274 100% activation (values equal to 1 on the y-axis) efficiency.
1275
1276 Quantification of the relative repressing and activating efficiencies of different Antp
1277 isoforms
1278 Quantification of the relative efficiency of Antp activating and repressing
1279 isoforms (Figure 4 D and H) were performed in Fiji (Schindelin et al., 2012) by
1280 selection of the total region of repression or activation of Antp protein or P1 reporter
1281 staining and quantification of the relative fluorescence intensity of the selected
1282 regions. 5-7 imaginal disc images per investigated genotype were used for analysis.
1283 For the repression assay the obtained values have been normalized over the
1284 intensity of Antp protein calculated in the region of overlap between eGFP and Antp
1285 (negative control). In both cases (repression and activation), the highest efficiency
1286 per transcript variant (for repression, the short linker isoform; for activation the long
1287 linker isoform) have been set to 100%.
1288
1289 Gel-shifts (Electrophoretic Mobility Shift Assays – EMSAs)
1290 Full-length Antp short and long linker variants (transcript variants RM and
1291 RN), encoding activating and repressing Antp isoforms, respectively, were cloned
1292 into the pET21b(+) vector (Novagen), which features a C-terminal 6xHis tag, and
1293 expressed in Rosetta™ 2 cells (Novagen), following the manufacturer's standard
1294 protocol. The two proteins were then Ni-column purified and subjected to gel-
1295 filtration. The concentrations of purified proteins were then compared by Western
1296 blotting, using the anti-Antp 4C3 antibody (Developmental Studies Hybridoma Bank,
1297 University of Iowa), and equal starting concentrations were used in successive

1298 twofold serial dilutions in gel-shift experiments. The BS1and BS2 binding sites have
1299 been identified ~2 kb upstream of the *engrailed* gene promoter and characterized for
1300 Antp binding previously (Affolter et al., 1990). The HB1 binding site has been
1301 described previously (Keegan et al., 1997) and is a binding site found in the intron of
1302 the mouse *Hoxa-4* gene. The D4 probe has been characterized previously (Duncan
1303 et al., 2010) as a functional element in the *spineless* gene.
1304
1305 Fluorescence Microscopy Imaging of live imaginal discs and FCS
1306 Fluorescence imaging and FCS measurements were performed on two
1307 uniquely modified confocal laser scanning microscopy systems, both comprised of
1308 the ConfoCor3 system (Carl Zeiss, Jena, Germany) and consisting of either an
1309 inverted microscope for transmitted light and epifluorescence (Axiovert 200 M); a
1310 VIS-laser module comprising the Ar/ArKr (458, 477, 488 and 514 nm), HeNe 543 nm
1311 and HeNe 633 nm lasers and the scanning module LSM510 META; or a Zeiss
1312 LSM780 inverted setup, comprising Diode 405 nm, Ar multiline 458, 488 and 514
1313 nm, DPSS 561 nm and HeNe 633 nm lasers. Both instruments were modified to
1314 enable detection using silicon Avalanche Photo Detectors (SPCM-AQR-1X;
1315 PerkinElmer, USA) for imaging and FCS. Images were recorded at a 512X512 pixel
1316 resolution. C-Apochromat 40x/1.2 W UV-VIS-IR objectives were used throughout.
1317 Fluorescence intensity fluctuations were recorded in arrays of 10 consecutive
1318 measurements, each measurement lasting 10 s. Averaged curves were analyzed
1319 using the software for online data analysis or exported and fitted offline using the
1320 OriginPro 8 data analysis software (OriginLab Corporation, Northampton, MA). In
1321 either case, the nonlinear least square fitting of the autocorrelation curve was
1322 performed using the Levenberg–Marquardt algorithm. Quality of the fitting was

1323 evaluated by visual inspection and by residuals analysis. Control FCS
1324 measurements to asses the detection volume were routinely performed prior to data
1325 acquisition, using dilute solutions of known concentration of Rhodamine 6G and
1326 Alexa488 dyes. The variability between independent measurements reflects
1327 variabilitys between cells, rather than imprecision of FCS measurements. For more
1328 details on Fluorescence Microscopy Imaging and FCS, refer to Supplement 1.
1329

1330 Sample size, biological and technical replicates

1331 For the measurement of TF molecular numbers and variability (Figure 1 and
1332 Figure 1 – figure supplement 1), 7-10 larvae of each fly strain were dissected,
1333 yielding at least 15 imaginal discs, which were used in FCS analysis. For the Fkh TF,
1334 7 pairs of salivary glands were analyzed and for AbdB, 12 genital discs were
1335 dissected from 12 larvae. More than 50 FCS measurements were performed in
1336 patches of neighboring cells of these dissected discs, in the regions of expression
1337 indicated in Figure 1 by arrows. Imaginal discs from the same fly strain (expressing a
1338 given endogenously-tagged TF) were analyzed on at least 3 independent instances
1339 (FCS sessions), taking place on different days (biological replicates) and for Antp,
1340 which was further analyzed in this study, more than 20 independent FCS sessions
1341 were used. As routinely done with FCS measurements in live cells, these
1342 measurements were evaluated during acquisition and subsequent analysis and,
1343 based on their quality (high counts per molecule and second, low photobleaching),
1344 were included in the calculation of concentration and variability. In Figure 1 – figure
1345 supplement 1 Q, *n* denotes the number of FCS measurements included in the
1346 calculations.

1347 For experiments involving immunostainings in imaginal discs to investigate
1348 the auto-regulatory behavior of Antp (Figures 2-5 and supplements thereof, except
1349 for the temporally-resolved auto-activating and repressing study of Antp in Figure 3,
1350 as discussed above), 14-20 male and female flies were mated in bottles and 10
1351 larvae were selected by means of fluorescent balancers and processed downstream.
1352 Up to 20 imaginal discs were visualized by fluorescence microscopy and high
1353 resolution Z-stacks were acquired for 3-5 representative discs or disc regions of
1354 interest per experiment. All experiments were performed in triplicate, except for the
1355 temporal analysis of Antp auto-regulatory behavior in Figure 3 (and figure
1356 supplements thereof), which was performed 6 times and the quantification of
1357 repression efficiency of short and long linker Antp isoforms in Figure 5 (and figure
1358 supplements thereof), which was performed 5 times.

1359 For the quantification of transcript variant abundance in Figure 4 D and H,
1360 RNA and thus cDNA was prepared from each stage 3 independent times (biological
1361 replicates) and the transcript abundance per RNA/cDNA sample was also analyzed
1362 3 times.

1363 For the experiments involving perturbations in Antp expression whereby the
1364 proper development of the leg and the notum have been assessed in Figure 5, more
1365 than 100 adult flies have been analyzed and this experiment has been performed
1366 more than 10 times independently.

1367

1368 **References**

1369 Abbott, M.K., and Kaufman, T.C. (1986). The relationship between the functional
1370 complexity and the molecular organization of the Antennapedia locus of Drosophila
1371 melanogaster. Genetics *114*, 919-942.

1372 Affolter, M., Percival-Smith, A., Muller, M., Leupin, W., and Gehring, W.J. (1990).
1373 DNA binding properties of the purified Antennapedia homeodomain. Proc Natl Acad
1374 Sci U S A *87*, 4093-4097.
1375 Becskei, A., and Serrano, L. (2000). Engineering stability in gene networks by
1376 autoregulation. Nature *405*, 590-593.
1377 Blake, W.J., M, K.A., Cantor, C.R., and Collins, J.J. (2003). Noise in eukaryotic gene
1378 expression. Nature *422*, 633-637.
1379 Boettiger, A.N., and Levine, M. (2013). Rapid transcription fosters coordinate snail
1380 expression in the Drosophila embryo. Cell reports *3*, 8-15.
1381 Buszczak, M., Paterno, S., Lighthouse, D., Bachman, J., Planck, J., Owen, S.,
1382 Skora, A.D., Nystul, T.G., Ohlstein, B., Allen, A.*, et al.* (2007). The carnegie protein
1383 trap library: a versatile tool for Drosophila developmental studies. Genetics *175*,
1384 1505-1531.
1385 Calleja, M., Moreno, E., Pelaz, S., and Morata, G. (1996). Visualization of gene
1386 expression in living adult Drosophila. Science *274*, 252-255.
1387 Clark, N.M., Hinde, E., Winter, C.M., Fisher, A.P., Crosti, G., Blilou, I., Gratton, E.,
1388 Benfey, P.N., and Sozzani, R. (2016). Tracking transcription factor mobility and
1389 interaction in Arabidopsis roots with fluorescence correlation spectroscopy. eLife *5*.
1390 Crickmore, M.A., Ranade, V., and Mann, R.S. (2009). Regulation of Ubx expression
1391 by epigenetic enhancer silencing in response to Ubx levels and genetic variation.
1392 PLoS genetics *5*, e1000633.
1393 Crocker, J., Abe, N., Rinaldi, L., McGregor, A.P., Frankel, N., Wang, S., Alsawadi,
1394 A., Valenti, P., Plaza, S., Payre, F.*, et al.* (2015). Low affinity binding site clusters
1395 confer hox specificity and regulatory robustness. Cell *160*, 191-203.

1396 Crocker, J., Noon, E.P., and Stern, D.L. (2016). The Soft Touch: Low-Affinity
1397 Transcription Factor Binding Sites in Development and Evolution. Current topics in
1398 developmental biology *117*, 455-469.
1399 Crocker, J., Tsai, A., Muthusamy, A.K., Lavis, L.K., Singer, R.H., and Stern, D.L.
1400 (2017). Nuclear microenvironments modulate transcription from low-affinity
1401 enhancers. bioRxiv.
1402 Dublanche, Y., Michalodimitrakis, K., Kummerer, N., Foglierini, M., and Serrano, L.
1403 (2006). Noise in transcription negative feedback loops: simulation and experimental
1404 analysis. Mol Syst Biol *2*, 41.
1405 Duncan, D., Kiefel, P., and Duncan, I. (2010). Control of the spineless antennal
1406 enhancer: direct repression of antennal target genes by Antennapedia. Dev Biol *347*,
1407 82-91.
1408 Dworkin, I., Lee, W., McCloskey, F., and Larsen, E. (2007). Complex genetic
1409 interactions govern the temporal effects of Antennapedia on antenna-to-leg
1410 transformations in Drosophila melanogaster. J Genet *86*, 111-123.
1411 Ejsmont, R.K., Bogdanzaliewa, M., Lipinski, K.A., and Tomancak, P. (2011).
1412 Production of fosmid genomic libraries optimized for liquid culture recombineering
1413 and cross-species transgenesis. Methods in molecular biology *772*, 423-443.
1414 Ejsmont, R.K., Sarov, M., Winkler, S., Lipinski, K.A., and Tomancak, P. (2009). A
1415 toolkit for high-throughput, cross-species gene engineering in Drosophila. Nat
1416 Methods *6*, 435-437.
1417 Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene
1418 expression in a single cell. Science *297*, 1183-1186.
1419 Engstrom, Y., Schneuwly, S., and Gehring, W.J. (1992). Spatial and Temporal
1420 Expression of an Antennapedia Lac Z Gene Construct Integrated into the

1421 Endogenous Antennapedia Gene of Drosophila-Melanogaster. Dev Genes Evol *201*,
1422 65-80.
1423 Franz, K., Singh, A., and Weinberger, L.S. (2011). Lentiviral vectors to study
1424 stochastic noise in gene expression. Methods in enzymology *497*, 603-622.
1425 Garaulet, D.L., Foronda, D., Calleja, M., and Sanchez-Herrero, E. (2008). Polycomb-
1426 dependent Ultrabithorax Hox gene silencing induced by high Ultrabithorax levels in
1427 Drosophila. Development *135*, 3219-3228.
1428 Gillespie, D.T. (2007). Stochastic simulation of chemical kinetics. Annu Rev Phys
1429 Chem *58*, 35-55.
1430 Gregor, T., Tank, D.W., Wieschaus, E.F., and Bialek, W. (2007a). Probing the limits
1431 to positional information. Cell *130*, 153-164.
1432 Gregor, T., Wieschaus, E.F., McGregor, A.P., Bialek, W., and Tank, D.W. (2007b).
1433 Stability and nuclear dynamics of the bicoid morphogen gradient. Cell *130*, 141-152.
1434 Gronlund, A., Lotstedt, P., and Elf, J. (2013). Transcription factor binding kinetics
1435 constrain noise suppression via negative feedback. Nat Commun *4*, 1864.
1436 Holloway, D.M., Lopes, F.J., da Fontoura Costa, L., Travencolo, B.A., Golyandina,
1437 N., Usevich, K., and Spirov, A.V. (2011). Gene expression noise in spatial patterning:
1438 hunchback promoter structure affects noise amplitude and distribution in Drosophila
1439 segmentation. PLoS computational biology *7*, e1001069.
1440 Holloway, D.M., and Spirov, A.V. (2015). Mid-embryo patterning and precision in
1441 Drosophila segmentation: Kruppel dual regulation of hunchback. PLoS One *10*,
1442 e0118450.
1443 Ito, K., Awano, W., Suzuki, K., Hiromi, Y., and Yamamoto, D. (1997). The Drosophila
1444 mushroom body is a quadruple structure of clonal units each of which contains a
1445 virtually identical set of neurones and glial cells. Development *124*, 761-771.

1446 Jorgensen, E.M., and Garber, R.L. (1987). Function and misfunction of the two
1447 promoters of the Drosophila Antennapedia gene. Genes Dev *1*, 544-555.
1448 Kang, M., Piliszek, A., Artus, J., and Hadjantonakis, A.K. (2013). FGF4 is required
1449 for lineage restriction and salt-and-pepper distribution of primitive endoderm factors
1450 but not their initial expression in the mouse. Development *140*, 267-279.
1451 Kaur, G., Costa, M.W., Nefzger, C.M., Silva, J., Fierro-Gonzalez, J.C., Polo, J.M.,
1452 Bell, T.D., and Plachta, N. (2013). Probing transcription factor diffusion dynamics in
1453 the living mammalian embryo with photoactivatable fluorescence correlation
1454 spectroscopy. Nat Commun *4*, 1637.
1455 Keegan, L.P., Haerry, T.E., Crotty, D.A., Packer, A.I., Wolgemuth, D.J., and Gehring,
1456 W.J. (1997). A sequence conserved in vertebrate Hox gene introns functions as an
1457 enhancer regulated by posterior homeotic genes in Drosophila imaginal discs.
1458 Mechanisms of development *63*, 145-157.
1459 Kelso, R.J., Buszczak, M., Quinones, A.T., Castiblanco, C., Mazzalupo, S., and
1460 Cooley, L. (2004). Flytrap, a database documenting a GFP protein-trap insertion
1461 screen in Drosophila melanogaster. Nucleic Acids Res *32*, D418-420.
1462 Krasnow, M.A., Saffman, E.E., Kornfeld, K., and Hogness, D.S. (1989).
1463 Transcriptional activation and repression by Ultrabithorax proteins in cultured
1464 Drosophila cells. Cell *57*, 1031-1043.
1465 Kuziora, M.A., and McGinnis, W. (1988). Autoregulation of a Drosophila homeotic
1466 selector gene. Cell *55*, 477-485.
1467 Lam, C.S., Mistri, T.K., Foo, Y.H., Sudhaharan, T., Gan, H.T., Rodda, D., Lim, L.H.,
1468 Chou, C., Robson, P., Wohland, T.*, et al.* (2012). DNA-dependent Oct4-Sox2
1469 interaction and diffusion properties characteristic of the pluripotent cell state revealed
1470 by fluorescence spectroscopy. The Biochemical journal *448*, 21-33.

1471 Le, T., Liang, Z., Patel, H., Yu, M.H., Sivasubramaniam, G., Slovitt, M., Tanentzapf,
1472 G., Mohanty, N., Paul, S.M., Wu, V.M., *et al.* (2006). A new family of Drosophila
1473 balancer chromosomes with a w- dfd-GMR yellow fluorescent protein marker.
1474 Genetics *174*, 2255-2257.
1475 LeMotte, P.K., Kuroiwa, A., Fessler, L.I., and Gehring, W.J. (1989). The homeotic
1476 gene Sex Combs Reduced of Drosophila: gene structure and embryonic expression.
1477 EMBO J *8*, 219-227.
1478 Li, Y., Wang, F., Lee, J.A., and Gao, F.B. (2006). MicroRNA-9a ensures the precise
1479 specification of sensory organ precursors in Drosophila. Genes Dev *20*, 2793-2805.
1480 Little, S.C., Tikhonov, M., and Gregor, T. (2013). Precise developmental gene
1481 expression arises from globally stochastic transcriptional activity. Cell *154*, 789-800.
1482 Lo, C.A., Kays, I., Emran, F., Lin, T.J., Cvetkovska, V., and Chen, B.E. (2015).
1483 Quantification of Protein Levels in Single Living Cells. Cell reports *13*, 2634-2644.
1484 Lucas, T., Ferraro, T., Roelens, B., De Las Heras Chanes, J., Walczak, A.M.,
1485 Coppey, M., and Dostatni, N. (2013). Live imaging of bicoid-dependent transcription
1486 in Drosophila embryos. Current biology : CB *23*, 2135-2139.
1487 Mistri, T.K., Devasia, A.G., Chu, L.T., Ng, W.P., Halbritter, F., Colby, D., Martynoga,
1488 B., Tomlinson, S.R., Chambers, I., Robson, P., *et al.* (2015). Selective influence of
1489 Sox2 on POU transcription factor binding in embryonic and neural stem cells. EMBO
1490 reports *16*, 1177-1191.
1491 Morin, X., Daneman, R., Zavortink, M., and Chia, W. (2001). A protein trap strategy
1492 to detect GFP-tagged proteins expressed from their endogenous loci in Drosophila.
1493 Proc Natl Acad Sci U S A *98*, 15050-15055.

1494 Muller, C.B., Loman, A., Pacheco, V., Koberling, F., Willbold, D., Richtering, W., and
1495 Enderlein, J. (2008). Precise measurement of diffusion by multi-color dual-focus
1496 fluorescence correlation spectroscopy. Epl-Europhys Lett *83*.
1497 Nevozhay, D., Adams, R.M., Murphy, K.F., Josic, K., and Balazsi, G. (2009).
1498 Negative autoregulation linearizes the dose-response and suppresses the
1499 heterogeneity of gene expression. Proc Natl Acad Sci U S A *106*, 5123-5128.
1500 Newman, J.R., Ghaemmaghami, S., Ihmels, J., Breslow, D.K., Noble, M., DeRisi,
1501 J.L., and Weissman, J.S. (2006). Single-cell proteomic analysis of S. cerevisiae
1502 reveals the architecture of biological noise. Nature *441*, 840-846.
1503 Papadopoulos, D.K., Krmpot, A.J., Nikolic, S.N., Krautz, R., Terenius, L., Tomancak,
1504 P., Rigler, R., Gehring, W.J., and Vukojevic, V. (2015). Probing the kinetic landscape
1505 of Hox transcription factor-DNA binding in live cells by massively parallel
1506 Fluorescence Correlation Spectroscopy. Mechanisms of development *138 Pt 2*, 218-
1507 225.
1508 Papadopoulos, D.K., Resendez-Perez, D., Cardenas-Chavez, D.L., Villanueva-
1509 Segura, K., Canales-del-Castillo, R., Felix, D.A., Funfschilling, R., and Gehring, W.J.
1510 (2011). Functional synthetic Antennapedia genes and the dual roles of YPWM motif
1511 and linker size in transcriptional activation and repression. Proc Natl Acad Sci U S A
1512 *108*, 11959-11964.
1513 Papadopoulos, D.K., Vukojevic, V., Adachi, Y., Terenius, L., Rigler, R., and Gehring,
1514 W.J. (2010). Function and specificity of synthetic Hox transcription factors in vivo.
1515 Proc Natl Acad Sci U S A *107*, 4087-4092.
1516 Pare, A., Lemons, D., Kosman, D., Beaver, W., Freund, Y., and McGinnis, W.
1517 (2009). Visualization of individual Scr mRNAs during Drosophila embryogenesis
1518 yields evidence for transcriptional bursting. Current biology : CB *19*, 2037-2042.

1519 Paulsson, J. (2004). Summing up the noise in gene networks. Nature *427*, 415-418.
1520 Pavlopoulos, A., and Akam, M. (2011). Hox gene Ultrabithorax regulates distinct sets
1521 of target genes at successive stages of Drosophila haltere morphogenesis. Proc Natl
1522 Acad Sci U S A *108*, 2855-2860.
1523 Pelaez, N., Gavalda-Miralles, A., Wang, B., Navarro, H.T., Gudjonson, H., Rebay, I.,
1524 Dinner, A.R., Katsaggelos, A.K., Amaral, L.A., and Carthew, R.W. (2015). Dynamics
1525 and heterogeneity of a fate determinant during transition towards cell differentiation.
1526 eLife *4*.
1527 Perez-Camps, M., Tian, J., Chng, S.C., Sem, K.P., Sudhaharan, T., Teh, C.,
1528 Wachsmuth, M., Korzh, V., Ahmed, S., and Reversade, B. (2016). Quantitative
1529 imaging reveals real-time Pou5f3-Nanog complexes driving dorsoventral
1530 mesendoderm patterning in zebrafish. eLife *5*.
1531 Plachta, N., Bollenbach, T., Pease, S., Fraser, S.E., and Pantazis, P. (2011). Oct4
1532 kinetics predict cell lineage patterning in the early mammalian embryo. Nat Cell Biol
1533 *13*, 117-123.
1534 Quinones-Coello, A.T., Petrella, L.N., Ayers, K., Melillo, A., Mazzalupo, S., Hudson,
1535 A.M., Wang, S., Castiblanco, C., Buszczak, M., Hoskins, R.A.*, et al.* (2007).
1536 Exploring strategies for protein trapping in Drosophila. Genetics *175*, 1089-1104.
1537 Rao, C.V., and Arkin, A.P. (2003). Stochastic chemical kinetics and the quasi-
1538 steady-state assumption: Application to the Gillespie algorithm. J Chem Phys *118*,
1539 4999-5010.
1540 Reed, H.C., Hoare, T., Thomsen, S., Weaver, T.A., White, R.A., Akam, M., and
1541 Alonso, C.R. (2010). Alternative splicing modulates Ubx protein function in
1542 Drosophila melanogaster. Genetics *184*, 745-758.

1543 Saadaoui, M., Merabet, S., Litim-Mecheri, I., Arbeille, E., Sambrani, N., Damen, W.,
1544 Brena, C., Pradel, J., and Graba, Y. (2011). Selection of distinct Hox-Extradenticle
1545 interaction modes fine-tunes Hox protein activity. Proc Natl Acad Sci U S A *108*,
1546 2276-2281.
1547 Sanchez, A., Garcia, H.G., Jones, D., Phillips, R., and Kondev, J. (2011). Effect of
1548 promoter architecture on the cell-to-cell variability in gene expression. PLoS
1549 computational biology *7*, e1001100.
1550 Sarov, M., Barz, C., Jambor, H., Hein, M.Y., Schmied, C., Suchold, D., Stender, B.,
1551 Janosch, S., K, J.V., Krishnan, R.T.*, et al.* (2016). A genome-wide resource for the
1552 analysis of protein localisation in Drosophila. eLife *5*, e12068.
1553 Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,
1554 Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B.*, et al.* (2012). Fiji: an open-source
1555 platform for biological-image analysis. Nat Methods *9*, 676-682.
1556 Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen,
1557 W., and Selbach, M. (2011). Global quantification of mammalian gene expression
1558 control. Nature *473*, 337-342.
1559 Sheth, R., Bastida, M.F., Kmita, M., and Ros, M. (2014). "Self-Regulation," A New
1560 Facet of Hox Genes' Function. Dev Dynam *243*, 182-191.
1561 Shimoga, V., White, J.T., Li, Y., Sontag, E., and Bleris, L. (2013). Synthetic
1562 mammalian transgene negative autoregulation. Mol Syst Biol *9*, 670.
1563 Slattery, M., Voutev, R., Ma, L., Negre, N., White, K.P., and Mann, R.S. (2013).
1564 Divergent transcriptional regulatory logic at the intersection of tissue growth and
1565 developmental patterning. PLoS genetics *9*, e1003753.
1566 Smolander, O.P., Kandhavelu, M., Mannerstrom, H., Lihavainen, E., Kalaichelvan,
1567 S., Healy, S., Yli-Harja, O., Karp, M., and Ribeiro, A.S. (2011). Cell-to-cell diversity in

1568 protein levels of a gene driven by a tetracycline inducible promoter. BMC molecular
1569 biology *12*, 21.
1570 Stroeher, V.L., Gaiser, J.C., and Garber, R.L. (1988). Alternative RNA splicing that is
1571 spatially regulated: generation of transcripts from the Antennapedia gene of
1572 Drosophila melanogaster with different protein-coding regions. Molecular and cellular
1573 biology *8*, 4143-4154.
1574 Swain, P.S., Elowitz, M.B., and Siggia, E.D. (2002). Intrinsic and extrinsic
1575 contributions to stochasticity in gene expression. Proc Natl Acad Sci U S A *99*,
1576 12795-12800.
1577 Szaloki, N., Krieger, J.W., Komaromi, I., Toth, K., and Vamosi, G. (2015). Evidence
1578 for Homodimerization of the c-Fos Transcription Factor in Live Cells Revealed by
1579 Fluorescence Microscopy and Computer Modeling. Molecular and cellular biology
1580 *35*, 3785-3798.
1581 Taniguchi, Y., Choi, P.J., Li, G.W., Chen, H., Babu, M., Hearn, J., Emili, A., and Xie,
1582 X.S. (2010). Quantifying E. coli proteome and transcriptome with single-molecule
1583 sensitivity in single cells. Science *329*, 533-538.
1584 Thattai, M., and van Oudenaarden, A. (2001). Intrinsic noise in gene regulatory
1585 networks. Proc Natl Acad Sci U S A *98*, 8614-8619.
1586 Tiwari, M., Mikuni, S., Muto, H., and Kinjo, M. (2013). Determination of dissociation
1587 constant of the NFkappaB p50/p65 heterodimer using fluorescence cross-correlation
1588 spectroscopy in the living cell. Biochemical and biophysical research
1589 communications *436*, 430-435.
1590 Tsutsumi, M., Muto, H., Myoba, S., Kimoto, M., Kitamura, A., Kamiya, M., Kikukawa,
1591 T., Takiya, S., Demura, M., Kawano, K.*, et al.* (2016). In vivo fluorescence correlation

1592 spectroscopy analyses of FMBP-1, a silkworm transcription factor. FEBS open bio *6*,
1593 106-125.
1594 Venken, K.J., Schulze, K.L., Haelterman, N.A., Pan, H., He, Y., Evans-Holm, M.,
1595 Carlson, J.W., Levis, R.W., Spradling, A.C., Hoskins, R.A.*, et al.* (2011). MiMIC: a
1596 highly versatile transposon insertion resource for engineering Drosophila
1597 melanogaster genes. Nat Methods *8*, 737-743.
1598 Vukojevic, V., Heidkamp, M., Ming, Y., Johansson, B., Terenius, L., and Rigler, R.
1599 (2008). Quantitative single-molecule imaging by confocal laser scanning microscopy.
1600 Proc Natl Acad Sci U S A *105*, 18176-18181.
1601 Vukojevic, V., Papadopoulos, D.K., Terenius, L., Gehring, W.J., and Rigler, R.
1602 (2010). Quantitative study of synthetic Hox transcription factor-DNA interactions in
1603 live cells. Proc Natl Acad Sci U S A *107*, 4093-4098.
1604 Zechner, C., Ruess, J., Krenn, P., Pelet, S., Peter, M., Lygeros, J., and Koeppl, H.
1605 (2012). Moment-based inference predicts bimodality in transient gene expression.
1606 Proc Natl Acad Sci U S A *109*, 8340-8345.
1607 Zink, B., Engstrom, Y., Gehring, W.J., and Paro, R. (1991). Direct interaction of the
1608 Polycomb protein with Antennapedia regulatory sequences in polytene
1609 chromosomes of Drosophila melanogaster. EMBO J *10*, 153-162.

1610

1611 **Supplemental information**

1612 Supplement 1: Background on Fluorescence Microscopy Imaging and FCS

1613 Two individually modified instruments (Zeiss, LSM 510 and 780, ConfoCor 3)
1614 with fully integrated FCS/CLSM optical pathways were used for imaging. The
1615 detection efficiency of CLSM imaging was significantly improved by the introduction
1616 of APD detectors. As compared to PMTs, which are normally used as detectors in

conventional CLSM, the APDs are characterized by higher quantum yield and collection efficiency – about 70 % in APDs as compared to 15 – 25 % in PMTs, higher gain, negligible dark current and better efficiency in the red part of the spectrum. Enhanced fluorescence detection efficiency enabled image collection using fast scanning ($1 - 5\ \mu s/pixel$). This enhances further the signal-to-noise-ratio by avoiding fluorescence loss due to triplet state formation, enabling fluorescence imaging with single-molecule sensitivity. In addition, low laser intensities ($150 - 750\ \mu W$) could be applied for imaging, significantly reducing the photo-toxicity (Vukojevic et al., 2008).

FCS measurements are performed by recording fluorescence intensity fluctuations in a very small, approximately ellipsoidal observation volume element (OVE) (about $0.2\mu m$ wide and $1\mu m$ long) that is generated in imaginal disc cells by focusing the laser light through the microscope objective and by collecting the fluorescence light through the same objective using a pinhole in front of the detector to block out-of-focus light. The fluorescence intensity fluctuations, caused by fluorescently labeled molecules passing through the OVE are analyzed using temporal autocorrelation analysis.

In temporal autocorrelation analysis we first derive the autocorrelation function $G(\tau)$:

$$G(\tau) = 1 + \frac{\langle \delta I(t) \cdot \delta I(t+\tau) \rangle}{\langle I(t) \rangle^2} \quad \text{(S1)},$$

where $\delta I(t) = I(t) - \langle I(t) \rangle$ is the deviation from the mean intensity at time t and $\delta I(t+\tau) = I(t+\tau) - \langle I(t) \rangle$ is the deviation from the mean intensity at time $t + \tau$. For further analysis, an autocorrelation curve is derived by plotting $G(\tau)$ as a function of the lag time, i.e. the autocorrelation time τ.

To derive information about molecular numbers and their corresponding diffusion time, the experimentally obtained autocorrelation curves are compared to autocorrelation functions derived for different model systems. A model describing free three dimensional (3D) diffusion of two components and triplet formation was used in this study:

$$G(\tau) = 1 + \frac{1}{N}\left(\frac{1-y}{\left(1+\frac{\tau}{\tau_{D_1}}\right)\cdot\sqrt{1+\frac{w_{xy}^2\tau}{w_z^2\tau_{D_1}}}} + \frac{y}{\left(1+\frac{\tau}{\tau_{D_2}}\right)\cdot\sqrt{1+\frac{w_{xy}^2\tau}{w_z^2\tau_{D_2}}}}\right)\cdot\left(1+\frac{T}{1-T}\cdot e^{-\frac{\tau}{\tau_T}}\right)$$

(S2)

In the above equation, *N* is the average number of molecules in the OVE; *y* is the fraction of the slowly moving Antp-eGFP molecules; τ_{D_1} is the diffusion time of the free Antp-eGFP molecules; τ_{D_2} is the diffusion time of Antp-eGFP molecules undergoing interactions with the DNA; w_{xy} and w_z are radial and axial parameters, respectively, related to spatial properties of the OVE; T is the average equilibrium fraction of molecules in the triplet state; and τ_T the triplet correlation time related to rate constants for intersystem crossing and the triplet decay. Spatial properties of the detection volume, represented by the square of the ratio of the axial and radial parameters ($\left(\frac{w_z}{w_{xy}}\right)^2$), are determined in calibration measurements performed using a solution of Rhodamine 6G for which the diffusion coefficient (D) is known to be $D_{Rh6G} = 4.1 \cdot 10^{-10}\, m^2 s^{-1}$ (Muller et al., 2008). The diffusion time, τ_D, measured by FCS, is related to the translation diffusion coefficient *D* by:

$$\tau_D = \frac{w_{xy}^2}{4D} \text{ (S3).}$$

To establish that Antp molecules diffusing through the OVE are the underlying cause of the recorded fluorescence intensity fluctuations, we plotted the characteristic decay times τ_{D1} and τ_{D2}, obtained by FCS, as a function of the total concentration of Antp molecules (Figure

1662 1 – figure supplement 2). We observed that both characteristic decay times remain stable for
1663 increasing total concentration of Antp molecules, signifying that the underlying process
1664 triggering the fluorescence intensity fluctuations is diffusion of fluorescent Antp molecules
1665 through the OVE (which should be independent of the total concentration of Antp
1666 molecules).

1667 Supplement 2: Calculation of the concentration of endogenous TFs and average
1668 number of molecules in imaginal disc cell nuclei from FCS measurements
1669 (exemplified for Antp)

1670 The Observation Volume Element (OVE), which can be approximated by a
1671 prolate ellipsoid, has a volume of:

1672 $V_{OVE} = \pi^{\frac{3}{2}} \cdot w_{xy}^2 \cdot z_0 = 5.57 \cdot 0.1847^{\mathbf{2}} \cdot 1 = 0.223 \cdot 10^{-18}\, m^3 = 0.19 \cdot 10^{-15} L$ (S4).

1673 Therefore, one fluorescent molecule in the OVE yields equal concentration to:

1674 $\frac{1}{0.19 \cdot 10^{-15}} = 5.263 \cdot 10^{15}$ molecules per L.

1675 $1\, M$ of Antp-eGFP molecules equals $6.022 \cdot 10^{23}$ molecules/L,

1676 $x\, M$ of Antp-eGFP molecules equal $5.263 \cdot 10^{15}$ molecules/L.

1677 Therefore:

1678 $x = \frac{5.263 \cdot 10^{15}}{6.022 \cdot 10^{23}} = 0.874 \cdot 10^{-8}\, M = 8.74\, nM$ (S5).

1679 This indicates that one molecule in the OVE corresponds on the average to
1680 $8.74\, nM$ concentration of fluorescent molecules in the nucleus.

1681 The wing disc cells within the Antp expression domain (prescutum precursors)
1682 are not spherical, but rather ellipsoidal. Their axes were determined by fluorescence
1683 imaging to be 1.4 μm in the transverse dimension and 2.8 μm in the longitudinal. The
1684 volume of the nucleus was approximated by the volume of a prolate ellipsoid:

$$V_{nucleus} = \frac{4}{3}\pi a^2 b = \frac{4}{3} \cdot 3.14 \cdot (1.4 \cdot 10^{-6})^2 \cdot 2.8 \cdot 10^{-6}\, m^3 = 22.99 \cdot 10^{-18}\, m^3 = 22.99 \cdot 10^{-15}\, L \text{ (S6).}$$

Therefore, the OVE represents roughly $1/121$ of the nuclear volume:

$$\frac{V_{nucleus}}{V_{OVE}} = 121 \text{ (S7)}$$

and the number of molecules in Antp-eGFP nuclei is on the average $57.37 \cdot 121 \approx 6942$ molecules.

The concentration of Antp in the wing disc nuclei is calculated as follows:

$1\, mol$ of Antp-eGFP molecules equals $6.023 \cdot 10^{23}$,

$n\, mol$ of Antp-eGFP molecules equal 6942.

Thus $n_{Antp-eGFP} = \frac{6942}{6.022 \cdot 10^{23}} = 1152.77 \cdot 10^{-23}\, mol$ (S8).

Thus the concentration of Antp-eGFP within the diploid wing disc nucleus will be:

$$C_{Antp-eGFP} = \frac{1182.77 \cdot 10^{-23}}{22.99 \cdot 10^{-15}} = 50.1 \cdot 10^{-8}\, M \approx 501\, nM \text{ (S9),}$$

equivalent to $57.37 \cdot 8.74 \approx 501\, nM$.

Fitting of autocorrelation curves using a model for free three dimensional (3D) diffusion of two components and triplet, revealed that on the average 63% of the total Antp-eGFP molecules are fast moving and appear to be freely diffusing in the nucleus, whereas 37% are slow and are likely participating in the formation of complexes with chromatin.

<u>Supplement 3: Calculation of the ratio of apparent Antp dissociation constant for short and long linker Antp isoforms from FCS measurements on ectopically expressed Antp</u>

The Antp TF undergoes both specific and non-specific interactions with DNA. Assuming that non-specific interactions precede the specific ones, a two-step process of consecutive reactions is anticipated (Vukojevic et al., 2010):

$$\begin{aligned} &DNA_{ns} + Antp - eGFP \underset{k_{-ns}}{\overset{k_{ns}}{\Longleftrightarrow}} (DNA - Antp - eGFP)_{ns} \\ &(DNA - Antp - eGFP)_{ns} + DNA_s \underset{k_{-s}}{\overset{k_s}{\Longleftrightarrow}} (DNA - Antp - eGFP)_s \end{aligned} \quad \text{(S10).}$$

The turnover rate for the non-specific complex is:

$$\frac{d[(DNA-Antp-eGFP)_{ns}]}{dt} = k_{ns} \cdot [DNA_{ns}] \cdot [Antp - eGFP] - (k_{-ns} + k_s \cdot [DNA_s]) \cdot [(DNA - Antp - eGFP)_{ns}] + k_{-s} \cdot [(DNA - Antp - eGFP)_s] \text{ (S11).}$$

Assuming a quasi-steady state approximation:

$$\frac{d[(DNA-Antp-eGFP)_{ns}]}{dt} = 0 \text{ (S11),}$$

$$(k_{-ns} + k_s \cdot [DNA_s]) \cdot [(DNA - Antp - eGFP)_{ns}] = k_{ns} \cdot [DNA_{ns}] \cdot [Antp - eGFP] + k_{-s} \cdot [(DNA - Antp - eGFP)_s] \text{ (S12).}$$

Using the mass balance equation to express the concentration of the free TF:

$$[Antp - eGFP] = [Antp - eGFP]_0 - [(DNA - Antp - eGFP)_{ns}] - [(DNA - Antp - eGFP)_s] \text{ (S13)}$$

and assuming that:

$$[DNA]_{ns} \approx [DNA]_0 \text{ (S14),}$$

equation (S4) becomes:

$$(k_{-ns} + k_s \cdot [DNA_s]) \cdot [(DNA - Antp - eGFP)_{ns}] = k_{ns} \cdot [DNA]_0 \cdot [Antp - eGFP]_0 - [(DNA - Antp - eGFP)_{ns} - [(DNA - Antp - eGFP)_s] + k_{-s} \cdot [(DNA - Antp - eGFP)_s] \text{ (S15),}$$

$$(k_{-ns} + k_s \cdot [DNA_s] + k_{ns} \cdot [DNA]_0) \cdot [(DNA - Antp - eGFP)_{ns}] = k_{ns} \cdot [DNA]_0 \cdot ([Antp - eGFP]_0 - [(DNA - Antp - eGFP)_s]) + k_{-s} \cdot [(DNA - Antp - eGFP)_s] \text{ (S16),}$$

$$(k_{-ns} + k_s \cdot [DNA_s] + k_{ns} \cdot [DNA]_0) \cdot [(DNA - Antp - eGFP)_{ns}] = k_{ns} \cdot [DNA]_0 \cdot [Antp - eGFP]_0 - (k_{ns} \cdot [DNA]_0 - k_{-s}) \cdot [(DNA - Antp - eGFP)_s] \text{ (S17)},$$

$$[(DNA - Antp - eGFP)_{ns}] = \frac{k_{ns} \cdot [DNA]_0}{k_{-ns} + k_s \cdot [DNA_s] + k_{ns} \cdot [DNA]_0} \cdot [Antp - eGFP]_0 - \frac{k_{ns} \cdot [DNA]_0 - k_{-s}}{k_{-ns} + k_s \cdot [DNA_s] + k_{ns} \cdot [DNA]_0} \cdot [(DNA - Antp - eGFP)_s] \text{ (S18)}.$$

According to equation (S18) and the FCS data presented in Figure 4 – figure supplement 2, the slope of the linear dependence for

a) the short linker Antp isoform gives:

$$\frac{k_{ns} \cdot [DNA]_0}{k_{-ns} + k_s \cdot [DNA_s] + k_{ns} \cdot [DNA]_0} = 0.34 \text{ (S19)}$$

and the intercept:

$$\frac{k_{ns} \cdot [DNA]_0 - k_{-s}}{k_{-ns} + k_s \cdot [DNA_s] + k_{ns} \cdot [DNA]_0} \cdot [(DNA - Antp - eGFP)_s] = 5.31\, nM \text{ (S20)}.$$

If k_{-s} is small compared to $k_{ns} \cdot [DNA]_0$ and can therefore be neglected, then:

$$0.34 \cdot [(DNA - Antp - eGFP)_s] = 5.31\, nM \text{ (S21)}.$$

Thus, the concentration of specific complex between Antp-eGFP and DNA in the wing disc cell nuclei can be estimated to be:

$$[(DNA - Antp - eGFP)_s] = 15.62\, nM \text{ (S22)}.$$

The average concentration of free-diffusing Antp-eGFP molecules is determined as follows:

$$[Antp - eGFP]_{free} = [Antp - eGFP]_0 - [(DNA - Antp - eGFP)_{ns}] - [(DNA - Antp - eGFP)_s] = [Antp - eGFP]_0 - (0.34 \cdot [Antp - eGFP]_0 - 5.31) - [(DNA - Antp - eGFP)_s] = 785.28 - 0.34 \cdot 785.28 + 5.31 - 15.62 = 507.97 nM \text{ (S23)}.$$

Using the experimentally determined concentration of specific DNA–Antp-eGFP complexes (equation (S22)), we could estimate the dissociation constant for the specific DNA–Antp-eGFP, as a function of the total concentration of specific Antp binding sites, to be:

$$K_{d,Antp}^{short\ linker} = \frac{[DNA_s]_{free} \cdot [Antp-eGFP]_{free}}{[(DNA-Antp-eGFP)_s]} = \frac{[DNA_s]_{free} \cdot 507.97}{15.62} \approx ([DNA_s]_{free} \cdot 32.52)\ nM$$

(S24).

b) The long linker Antp isoform gives:

$$\frac{k_{ns} \cdot [DNA]_0}{k_{-ns} + k_s \cdot [DNA_s] + k_{ns} \cdot [DNA]_0} = 0.24 \text{ (S25).}$$

and the intercept:

$$\frac{k_{ns} \cdot [DNA]_0 - k_{-s}}{k_{-ns} + k_s \cdot [DNA_s] + k_{ns} \cdot [DNA]_0} \cdot [(DNA - Antp - eGFP)_s] = 3.28\ nM \text{ (S26).}$$

If k_{-s} is small compared to $k_{ns} \cdot [DNA]_0$ and can therefore be neglected, then:

$$0.24 \cdot [(DNA - Antp - eGFP)_s] = 3.28\ nM \text{ (S27).}$$

Thus, the concentration of specific complex between Antp-eGFP and DNA in the wing disc cell nuclei can be estimated to be:

$$[(DNA - Antp - eGFP)_s] = 13.67\ nM \text{ (S28).}$$

The average concentration of free-diffusing Antp-eGFP molecules is determined as follows:

$$[Antp - eGFP]_{free} = [Antp - eGFP]_0 - [(DNA - Antp - eGFP)_{ns}] - [(DNA - Antp - eGFP)_s] = [Antp - eGFP]_0 - (0.24 \cdot [Antp - eGFP]_0 - 3.28) - [(DNA - Antp - eGFP)_s] = 1382.95 - 0.24 \cdot 785.28 + 3.28 - 13.67 = 1040.65 nM \text{ (S29).}$$

Using the experimentally determined concentration of specific DNA–Antp-eGFP complexes (equation (S14)), we could estimate the dissociation constant for

the specific DNA–Antp-eGFP, as a function of the total concentration of specific Antp binding sites, to be:

$$K_{d,Antp}^{long\ linker} = \frac{[DNA_s]_{free} \cdot [Antp-eGFP]_{free}}{[(DNA-Antp-eGFP)_s]} = \frac{[DNA_s]_{free} \cdot 1040.65}{13.67} \approx ([DNA_s]_{free} \cdot 76.13)\ nM$$

(S30).

From equations (S24) and (S30), we could calculate the ratio of the apparent equilibrium dissociation constants to be:

$$\frac{K_{d,Antp}^{long\ linker}}{K_{d,Antp}^{short\ linker}} = \frac{[DNA_s]_{free}^{long\ linker} \cdot 76.13}{[DNA_s]_{free}^{short\ linker} \cdot 32.52} = 2.34 \cdot \frac{([DNA]_0 - [(DNA-Antp-eGFP)_s]_{long\ linker})}{([DNA]_0 - [(DNA-Antp-eGFP)_s]_{short\ linker})} =$$

$$2.34 \frac{[DNA]_0 - 13.67}{[DNA]_0 - 15.62}$$ (S31).

Therefore:

$$K_{d,Antp}^{long\ linker} > 2.34 \cdot K_{d,Antp}^{short\ linker}$$ (S32),

independently of the total concentration of Antp binding sites in the nucleus. For $[DNA]_0$ tending to $15.62\ nM$ ($[DNA]_0 \rightarrow 15.62\ nM$) with $[DNA]_0 > 15.62\ nM$, the ratio of the apparent equilibrium dissociation constants will be high:

$$\lim_{[DNA]_0 \rightarrow 15.62} \left(2.34 \frac{[DNA]_0 - 13.67}{[DNA]_0 - 15.62}\right) = +\infty$$ (S33),

indicating that the short linker isoform will bind the Antp binding sites with much higher affinity than the short linker isoform.

In contrast, for considerably higher values of $[DNA]_0$ ($[DNA]_0 \rightarrow +\infty$), the ratio of apparent equilibrium dissociation constants will be:

$$\lim_{[DNA]_0 \to +\infty} \left(2.34 \frac{[DNA]_0 - 13.67}{[DNA]_0 - 15.62}\right) =$$

$$\lim_{[DNA]_0 \to +\infty} \left(2.34 \frac{[DNA]_0\left(1 - \frac{13.67}{[DNA]_0}\right)}{[DNA]_0\left(1 - \frac{15.62}{[DNA]_0}\right)}\right) = \lim_{[DNA]_0 \to +\infty} \left(2.34 \frac{1 - \frac{13.67}{[DNA]_0}}{1 - \frac{15.62}{[DNA]_0}}\right) = 2.34 \text{ (S34)},$$

indicating a roughly 2.5fold higher affinity of the short linker repressive isoform. This means that, with decreasing total amounts of available Antp binding sites, the short linker repressing Antp isoform will bind with even higher affinity than its long linker counterpart.

Supplement 4: Stochastic modeling of Antennapedia expression

In the following, we develop a simple mathematical model that is able to explain the behavior of Antennapedia (Antp) expression at early and late developmental stages. The Antp promoter is modeled as a continuous-time Markov chain with three distinct transcriptional states. In the absence of Antp, the promoter is in an unbound state ("U"), in which transcription is inactive. From this state, the promoter can switch to a transcriptionally active state "A" at a rate, assumed to be proportional to the concentration of the long-linker, activating isoform of Antp. Analogously, repression of the promoter by the short-linker isoform of Antp is modeled by an additional transcriptionally weak state "R", which can be reached from state "U" at a rate proportional to the concentration of that isoform. The corresponding reverse transitions from states "R" and "A" back into state "U" are assumed to happen at a constant rate k. Since the activating isoform can potentially also repress the promoter, we assume that state "R" can be reached also from the active state "A". Similarly, we model a potential link also in the reverse direction from state "A" to "R". Depending on the model variant, we assume this transition happens either at a constant rate k (competitive promoter model) or at a rate proportional to

the concentration of the repressing isoform of Antp (non-competitive promoter model). In the latter case, repression through short-linker isoforms can take place even if a long-linker isoform is already bound to the promoter. As we have demonstrated in Figure 5 A and B, the two model variants yield qualitative differences in Antp expression. For the sake of illustration, the following description focuses on the non-competitive model variant but we remark that the competitive model can be derived analogously.

At a particular time point t, the transcription rate of Antp is determined by the current state of the promoter, i.e., $\lambda(t)\epsilon\{0, \lambda_A, \lambda_B\}$, with λ_A and λ_B as the transcription rates associated with states "A" and "R", respectively. In line with our experimental findings, we assume that transcripts are spliced into the activating and repressing isoforms at different rates ρ_A and ρ_R, respectively. This allows us to capture the imbalance between the two isoforms that was revealed by our FCS data. The overall expression rates for the two isoforms of Antp are then given by $h_A(t) = \lambda(t)Z\rho_A$ and $h_R(t) = \lambda(t)Z\rho_R$, whereas Z is a random variable that accounts for extrinsic variability in gene expression rates (Zechner et al., 2012). In all of our analyses, we model Z as a Gamma random variable, i.e., $Z \sim \Gamma(\alpha, \beta)$ with α and β as shape and inverse scale parameters of that distribution. In summary, we describe the auto-regulatory circuit of Antp expression by a stochastic reaction network of the form:

$$U \underset{k}{\overset{\gamma_A X_A(t)}{\rightleftharpoons}} A$$

$$U \underset{k}{\overset{\gamma_R X_R(t)}{\rightleftharpoons}} \mathrm{R}$$

$$A \underset{k}{\overset{\gamma_R X_R(t)}{\rightleftharpoons}} \mathrm{R}$$

$$\emptyset \xrightarrow{h_A(t)} X_A \xrightarrow{1/\tau} \emptyset$$

$$\emptyset \xrightarrow{h_R(t)} X_R \xrightarrow{1/\tau} \emptyset$$

(S35),

with $X_A(t)$and $X_R(t)$ as the concentration of the activating and repressing isoforms of Antp and τ as the protein half-live. The initial conditions $X_A(0)$ and $X_R(0)$ were drawn randomly in accordance with our concentration measurements at early stages. In particular, we assume that the total amount of Antp X_{tot} in each cell is drawn from a negative binomial distribution such that $X_{tot} \sim \mathcal{N}B(r,p)$, with r and p as the parameters characterizing this distribution. The total number of Antp molecules was then randomly partitioned into fractions of repressing and activating isoforms according to a Beta distribution. More specifically, we set $X_A(0) = WX_{tot}$ and $X_R(0) = (1-W)X_{tot}$ with $W \sim Beta(a,b)$. The parameters r, p, a and b were chosen based on our experimental data (see Table 1).

Due to the fact that Antp expression takes place at the timescale of hours, we can further simplify our model from (S35). In particular, we can make use of a quasi-steady state assumption (Rao and Arkin, 2003), taking into consideration that promoter switching due to binding and unbinding of the different Antp isoforms occurs at a much faster timescale than production and degradation of Antp. As a consequence, we can replace the stochastic gene expression rates of the two isoforms by their expected value, whereas the expectation is taken with respect to the quasi-stationary distribution of the three state promoter model. More precisely, we have:

$$h_A\,(t) \approx \mathbb{E}[\lambda(t)]Z\rho_A$$

$$h_R\,(t) \approx \mathbb{E}[\lambda(t)]Z\rho_R$$

(S36),

with $\mathbb{E}[\lambda(t)] = P_U 0 + P_A\lambda_A + P_R\lambda_R$ as the quasi-stationary probabilities of finding the promoter in state “U”, “A” and “R”, respectively. These probabilities can be derived from the generator matrix of the three-state promoter model, which reads:

$$Q = \begin{pmatrix} -\gamma_A X_A(t) - \gamma_R X_R(t) & k & k \\ \gamma_A X_A(t) & -k - \gamma_R X_R(t) & k \\ \gamma_R X_R(t) & \gamma_R X_R(t) & -2k \end{pmatrix} \text{(S37).}$$

Assuming that $X_A(t)$ and $X_R(t)$ remain roughly constant on the timescale of the promoter, the quasi-stationary distribution can be determined by the null-space of Q, which in this case is given by:

$$P_{QSS} = \begin{pmatrix} P_U \\ P_A \\ P_R \end{pmatrix} = \begin{pmatrix} \frac{k}{(k+\gamma_A X_A(t)+\gamma_R X_R(t))} \\ \frac{k(2\gamma_A X_A(t)+\gamma_R X_R(t))}{(2k+\gamma_R X_R(t))(k+\gamma_A X_A(t)+\gamma_R X_R(t))} \\ \frac{\gamma_R X_R(t)}{2k+\gamma_R X_R(t)} \end{pmatrix} \text{(S38).}$$

Correspondingly, the expectation of $\lambda(t)$ becomes:

$$\mathbb{E}[\lambda(t)] = (0 \quad \lambda_A \quad \lambda_R)\begin{pmatrix} P_U \\ P_A \\ P_R \end{pmatrix}$$

$$= \frac{\gamma_R X_R(t)\lambda_R}{2k + \gamma_R X_R(t)} + \frac{k\lambda_A\big(2\gamma_A X_A(t) + \gamma_R X_R(t)\big)}{\big(2k + \gamma_R X_R(t)\big)\big(k + \gamma_A X_A(t) + \gamma_R X_R(t)\big)}$$

$$:= \bar{\lambda}(X_A(t), X_R(t)) \text{ (S39).}$$

The simplified model of Antp expression can then be compactly written as:

$$\emptyset \xrightarrow{\bar{h}_A(X_A(t),X_R(t))} X_A \xrightarrow{1/\tau} \emptyset$$

$$\emptyset \xrightarrow{\bar{h}_R(X_A(t),X_R(t))} X_R \xrightarrow{1/\tau} \emptyset$$

(S40),

with $\bar{h}_A(X_A(t), X_R(t)) = \bar{\lambda}(X_A(t), X_R(t))Z\rho_A$ and $\bar{h}_R(X_A(t), X_R(t)) = \bar{\lambda}(X_A(t), X_R(t))Z\rho_R$.

In all our simulation studies, the circuit from (S40) was simulated using the stochastic simulation algorithm (SSA) (Gillespie, 2007). In case of the perturbation experiments, small modifications to the model were made. In case of overexpressing either of the two isoforms, additional reactions of the form:

$$\emptyset \xrightarrow{\lambda_O} X_A \text{ (S41)}$$

and

$$\emptyset \xrightarrow{\lambda_O} X_A \text{ (S42)}$$

were included in the model. To account for overexpression of an external repressor S, we introduced a fourth state in the promoter model, from which no expression can take place. This state is assumed to be reachable from any of the other three states at a rate $\gamma_S S$ with S as the concentration of the external repressor. The corresponding reaction rates of Antp expression were determined analogously to equations (S37-S39). Table 1 summarizes the parameters used for each of the simulation studies.

Table 1: Parameters used for simulating the stochastic model of Antp expression

Parameter	r	p	a	b	τ	k	γ_A	γ_R	λ_A	λ_R	α	β	λ_O	γ_S	S
Unit	$-$	$-$	$-$	$-$	h	s^{-1}	s^{-1}	s^{-1}	s^{-1}	s^{-1}	$-$	$-$	s^{-1}	$-$	$-$
Figure 5 A, D, E and F	2.5	1.0	18	12	12^a	100	0.25	1	10	0.1	3	0.33	$-$	$-$	$-$

Figure 5 B	2.5	1.0	18	12	12	100	0.25	1	10	0.1	3	0.33	–	–	–
Figure 5 C	2.5	1.0	18	12	12	100	0.25	0	0.5[b]	–	3	0.33	–	–	
Figure 5 G[c]	2.5	1.0	18	12	12	100	0.25	1	10	0.1	3	0.33	10	–	–
Figure 5 H[d]	2.5	1.0	18	12	12	100	0.25	1	10	0.1	3	0.33	10	–	–
Figure 5 M	2.5	1.0	18	12	12	100	0.25	1	10	0.1	3	0.33	10	1	$4e4$

[a]Experimentally determined value from (Dworkin et al., 2007).

[b]The expression rate λ_A was chosen to achieve average Antp levels similar to those of Figure 5 A (i.e., around 6000 molecules).

[c]Overexpression of X_R.

[d]Overexpression of X_A.

Figure 1

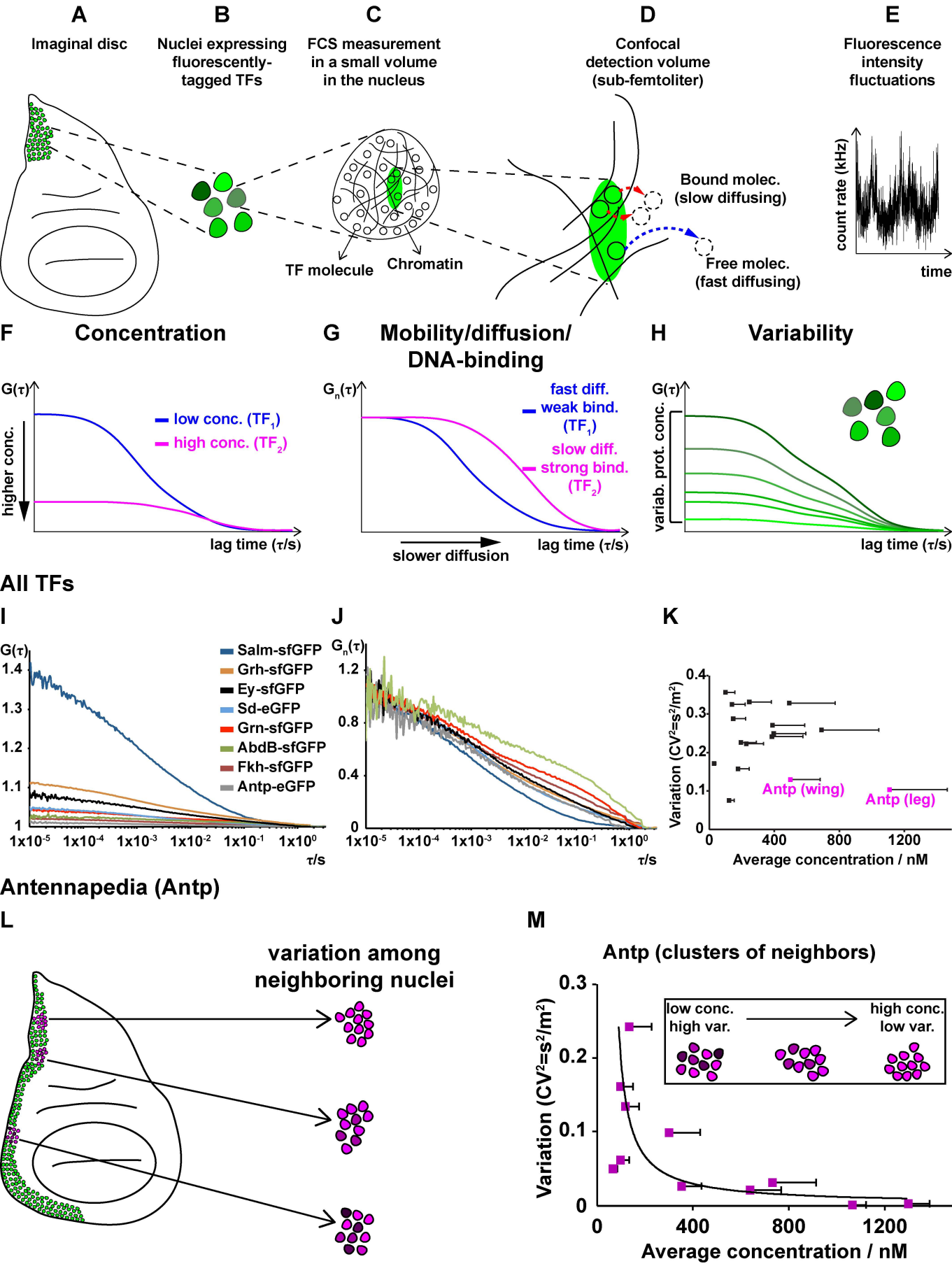
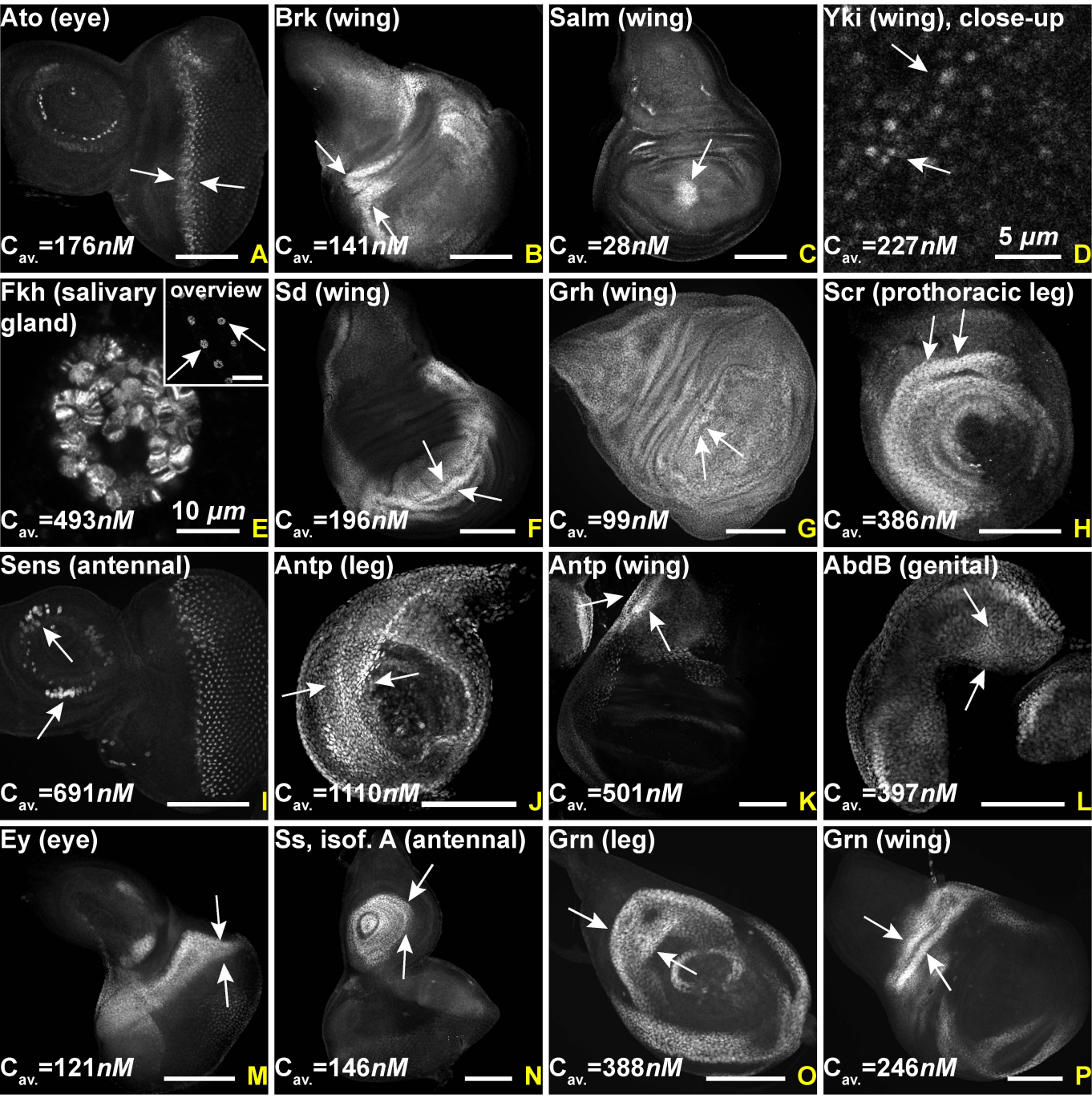

A
Imaginal disc
B
Nuclei expressing fluorescently-tagged TFs
C
FCS measurement in a small volume in the nucleus
D
Confocal detection volume (sub-femtoliter)
E
Fluorescence intensity fluctuations
TF molecule
Chromatin
Bound molec. (slow diffusing)
Free molec. (fast diffusing)
count rate (kHz)
time
F Concentration
G Mobility/diffusion/ DNA-binding
H Variability
G(τ)
higher conc.
low conc. (TF1)
high conc. (TF2)
lag time (τ/s)
Gn(τ)
fast diff.
weak bind. (TF1)
slow diff.
strong bind. (TF2)
slower diffusion
variab. prot. conc.
All TFs
I
J
K
Salm-sfGFP
Grh-sfGFP
Ey-sfGFP
Sd-eGFP
Grn-sfGFP
AbdB-sfGFP
Fkh-sfGFP
Antp-eGFP
τ/s
Variation (CV²=s²/m²)
Antp (wing)
Antp (leg)
Average concentration / nM
Antennapedia (Antp)
L
variation among neighboring nuclei
M
Antp (clusters of neighbors)
low conc. high var.
high conc. low var.

Figure 1 - figure supplement 1



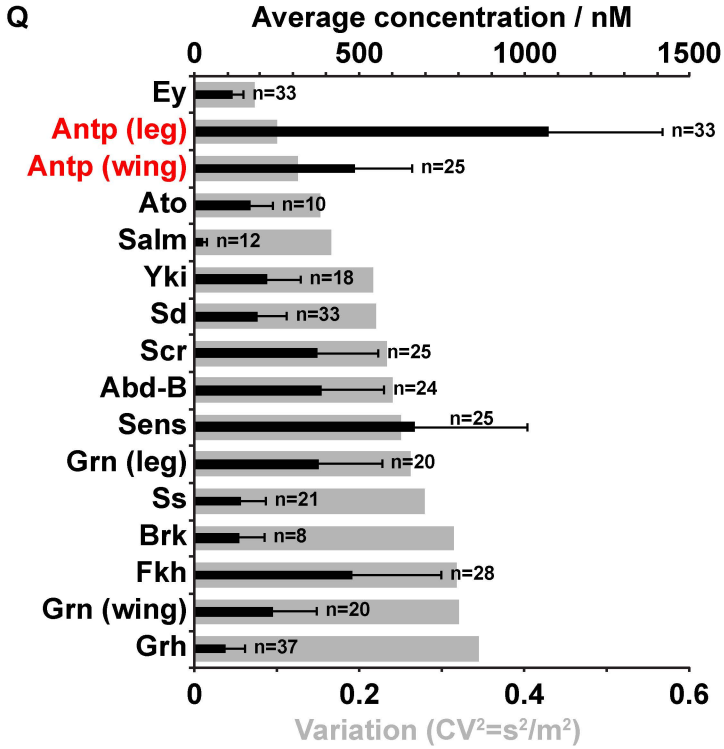


Figure 1 - figure supplement 2

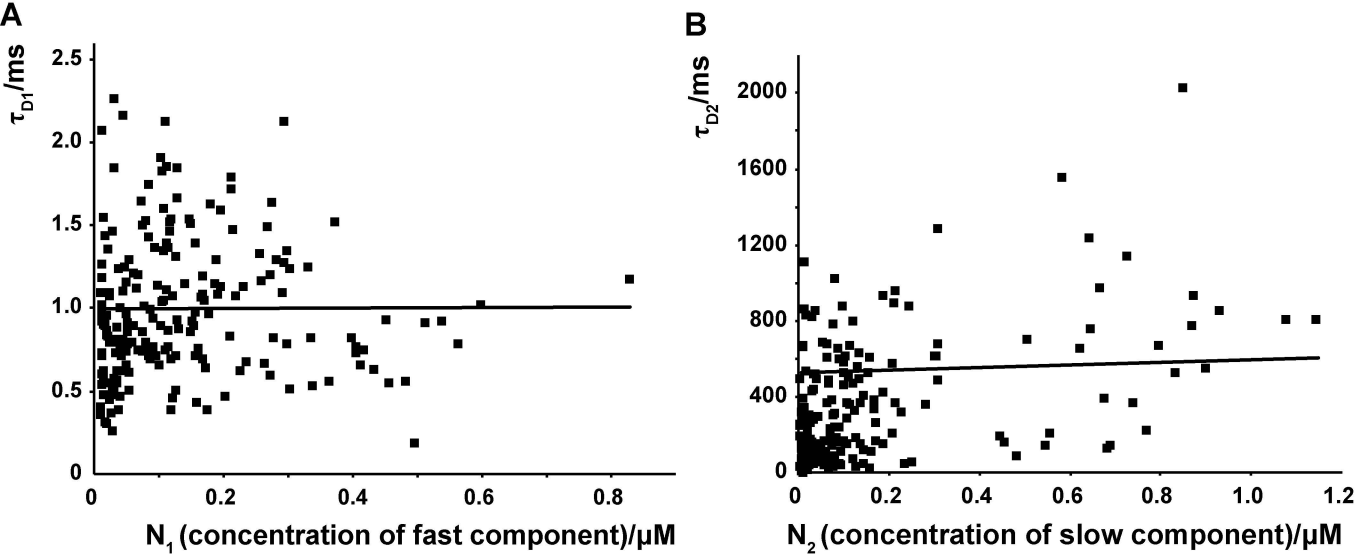
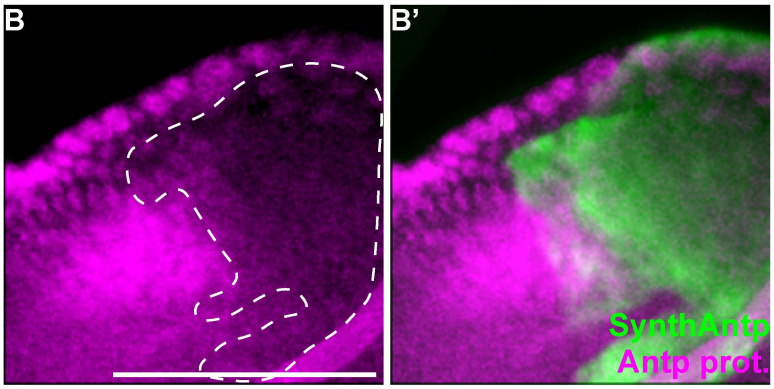


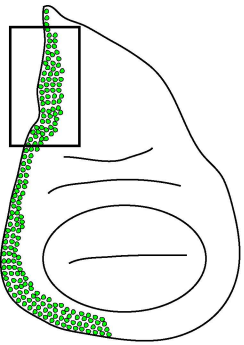
Figure 2

Auto-repression: protein level

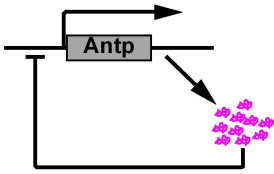
Ay>SynthAntp-eGFP



A

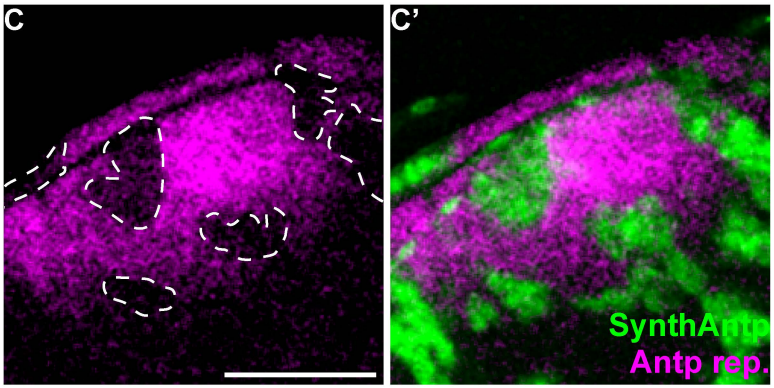


D



Auto-repression: transcripts level

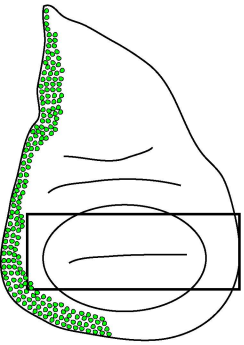
Antp (P1)-lacZ, Ay>SynthAntp-eGFP

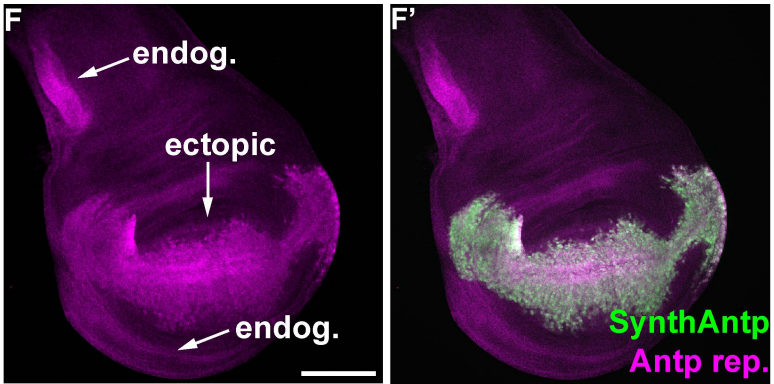


Auto-activation: transcripts level

Antp (P1)-lacZ, *Dll* (MD23)>SynthAntp-eGFP

E





G

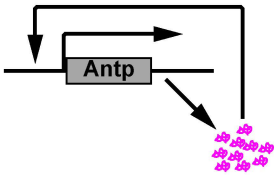


Figure 2 - figure supplement 1

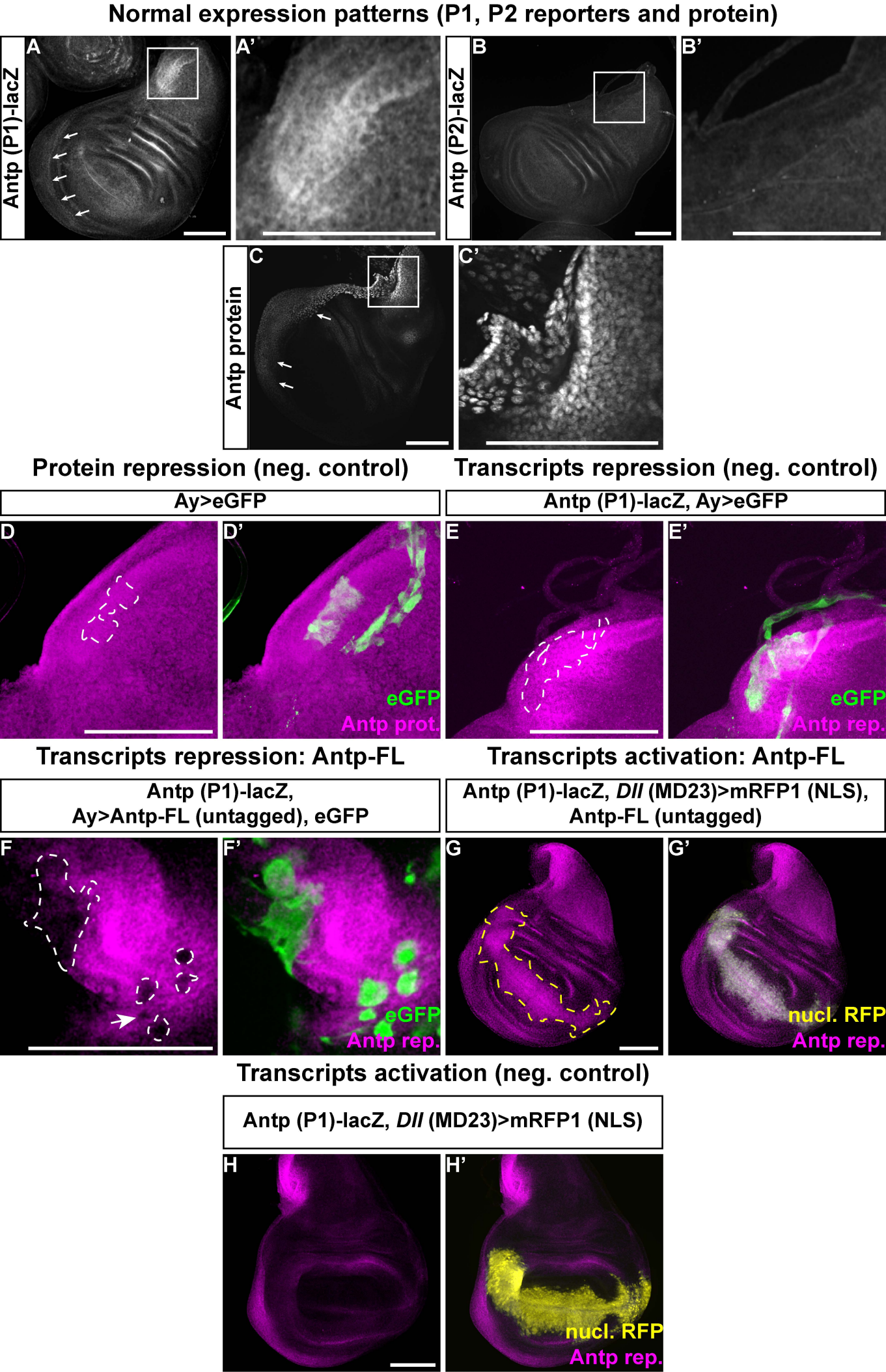
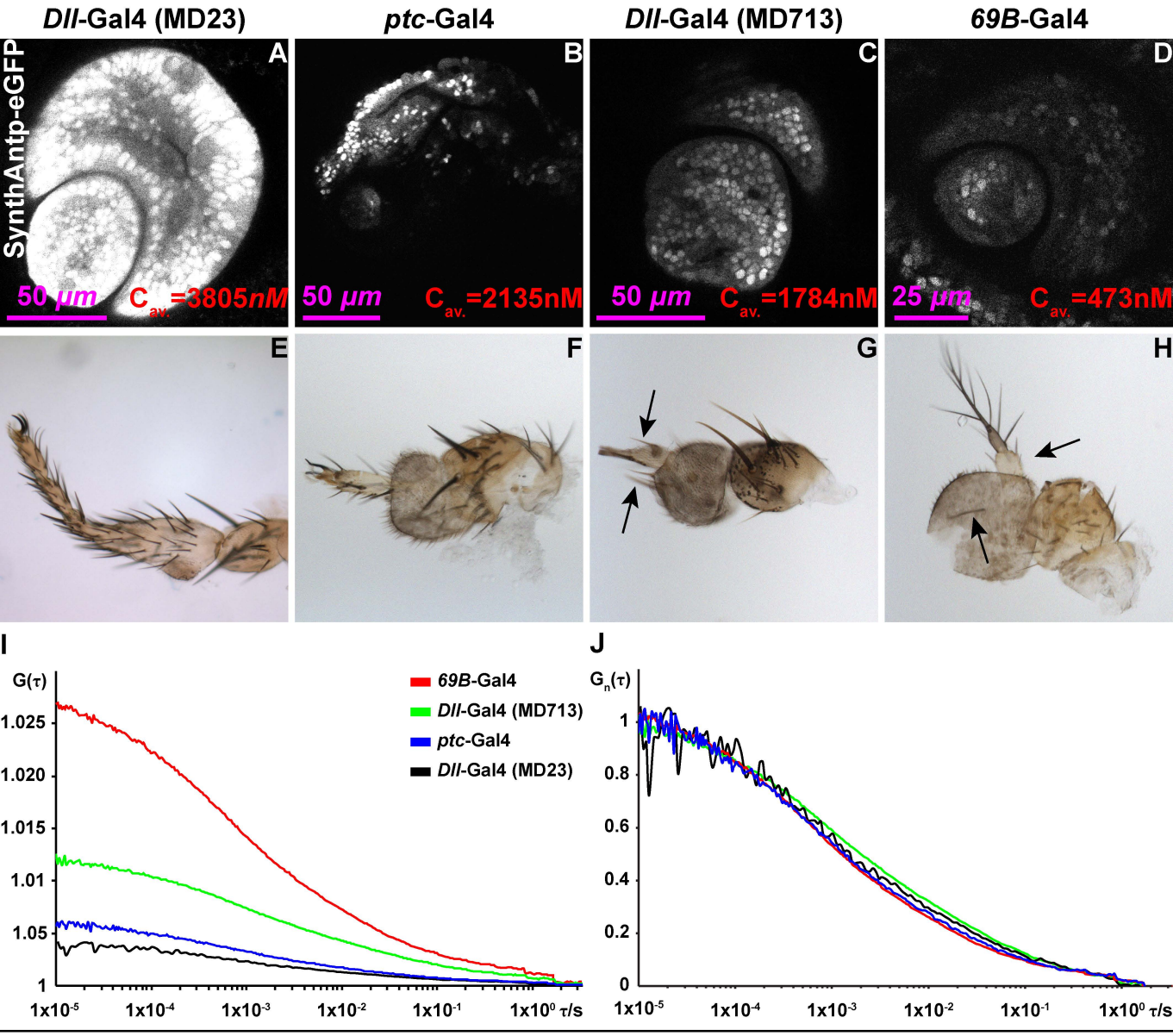
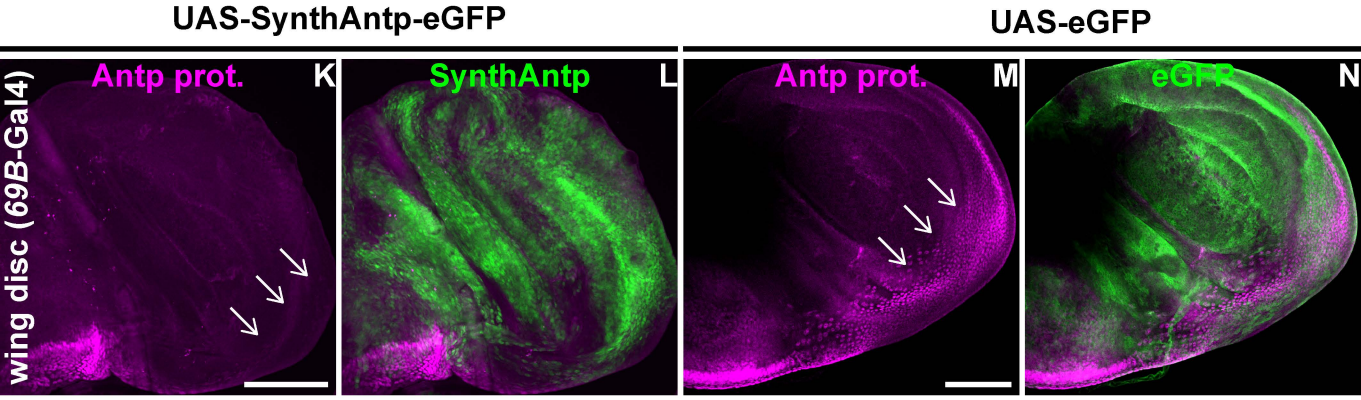


Figure 2 - figure supplement 2



Repression



Activation

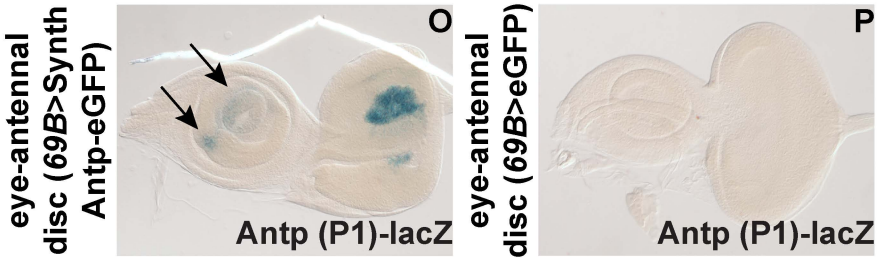


Figure 2 - figure supplement 3

Endogenous versus overexpressed

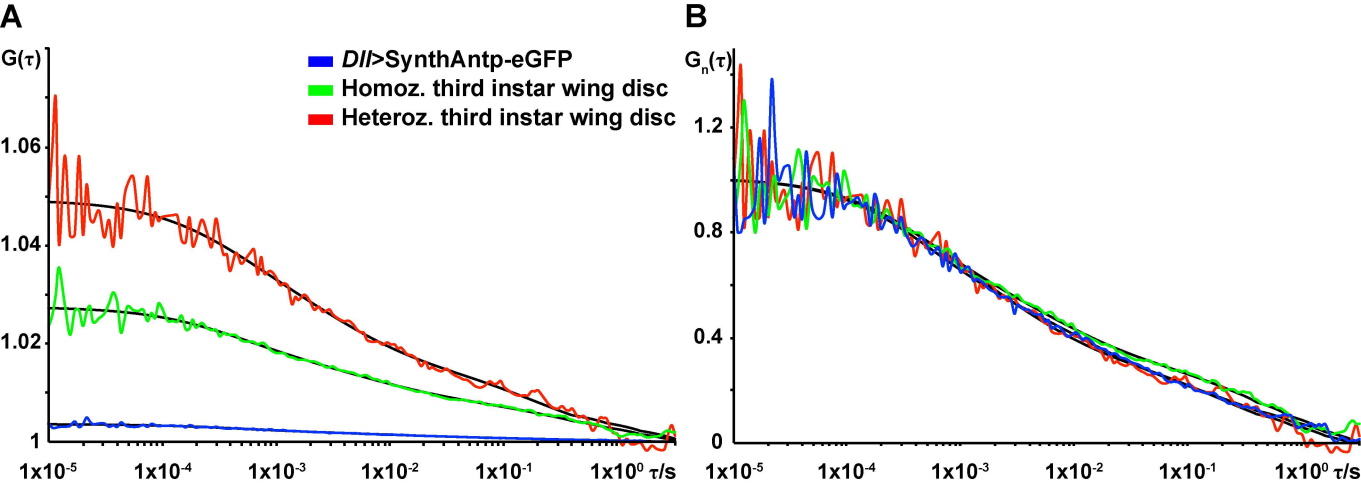


Figure 3

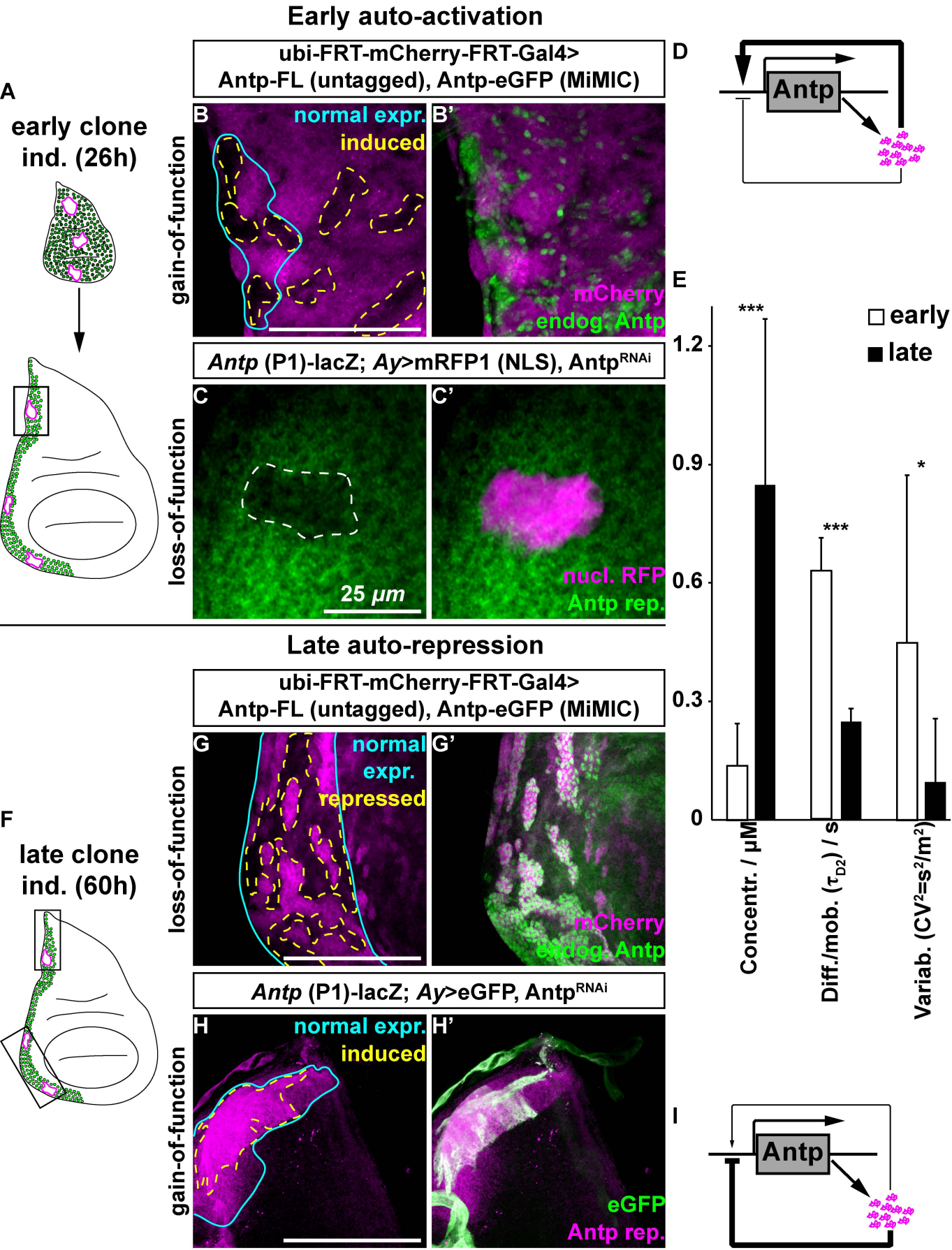

Early auto-activation
ubi-FRT-mCherry-FRT-Gal4>
Antp-FL (untagged), Antp-eGFP (MiMIC)
A
early clone
ind. (26h)
gain-of-function
B
normal expr.
induced
B'
mCherry
endog. Antp
Antp (P1)-lacZ; Ay>mRFP1 (NLS), Antp^RNAi
loss-of-function
C
25 μm
C'
nucl. RFP
Antp rep.
D
Antp
E
early
late
***
*
***
1.2
0.9
0.6
0.3
0
Concentr. / μM
Diff./mob. (τD2) / s
Variab. (CV²=s²/m²)
Late auto-repression
ubi-FRT-mCherry-FRT-Gal4>
Antp-FL (untagged), Antp-eGFP (MiMIC)
F
late clone
ind. (60h)
loss-of-function
G
normal
expr.
repressed
G'
mCherry
endog. Antp
Antp (P1)-lacZ; Ay>eGFP, Antp^RNAi
gain-of-function
H
normal expr.
induced
H'
eGFP
Antp rep.
I
Antp

Figure 3 - figure supplement 1

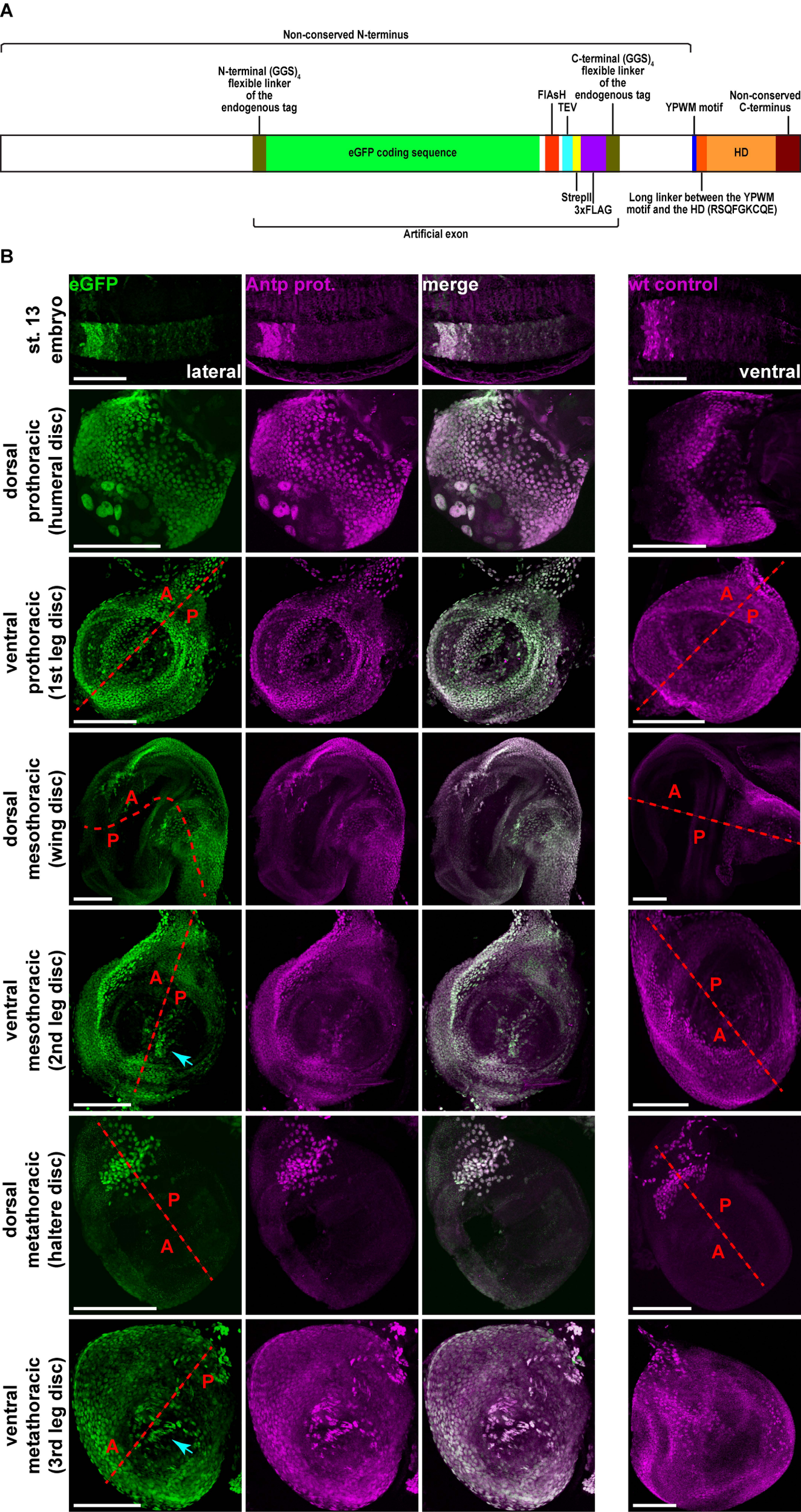


Figure 3 - figure supplement 2

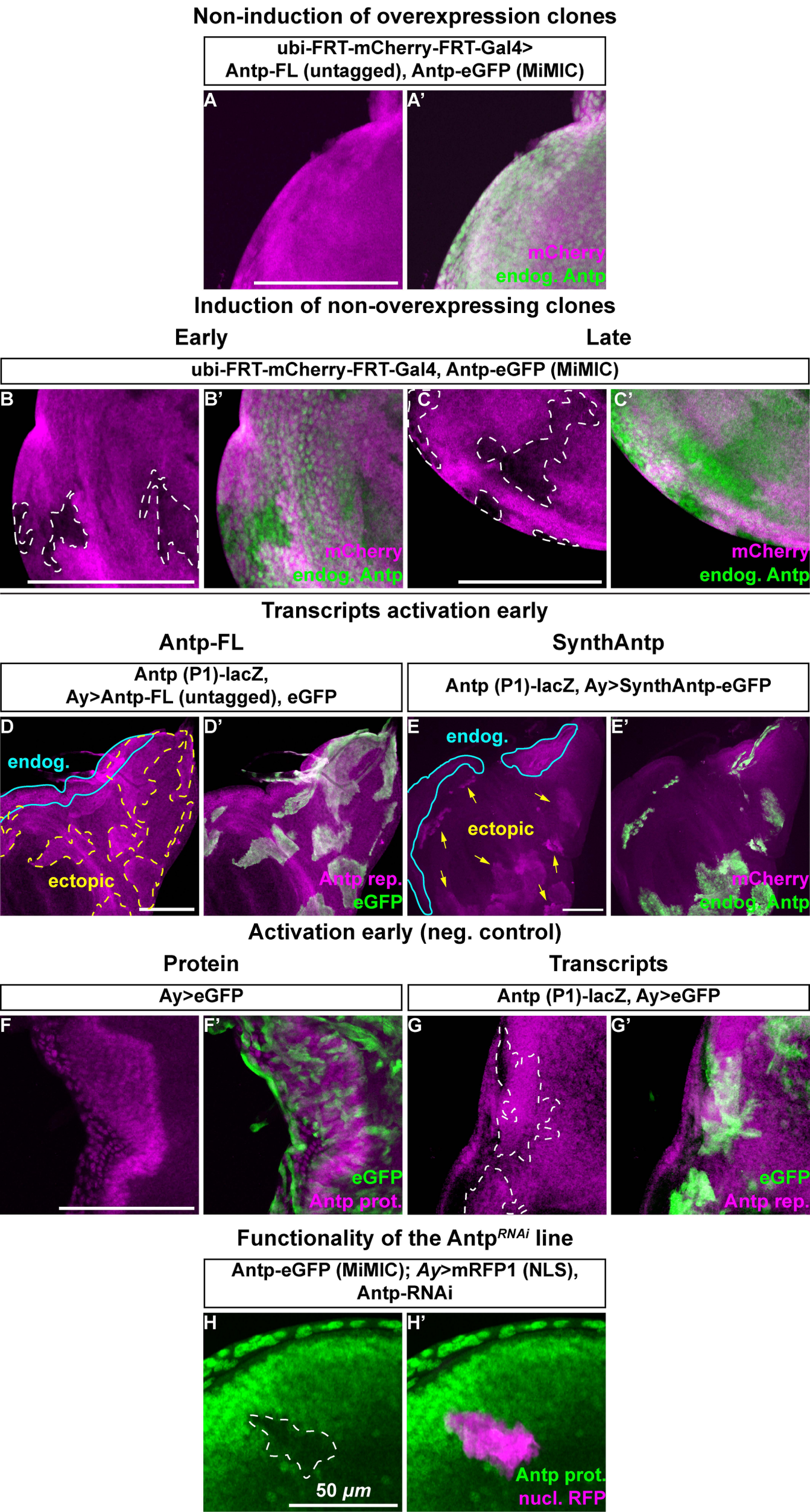


Figure 3 - figure supplement 3

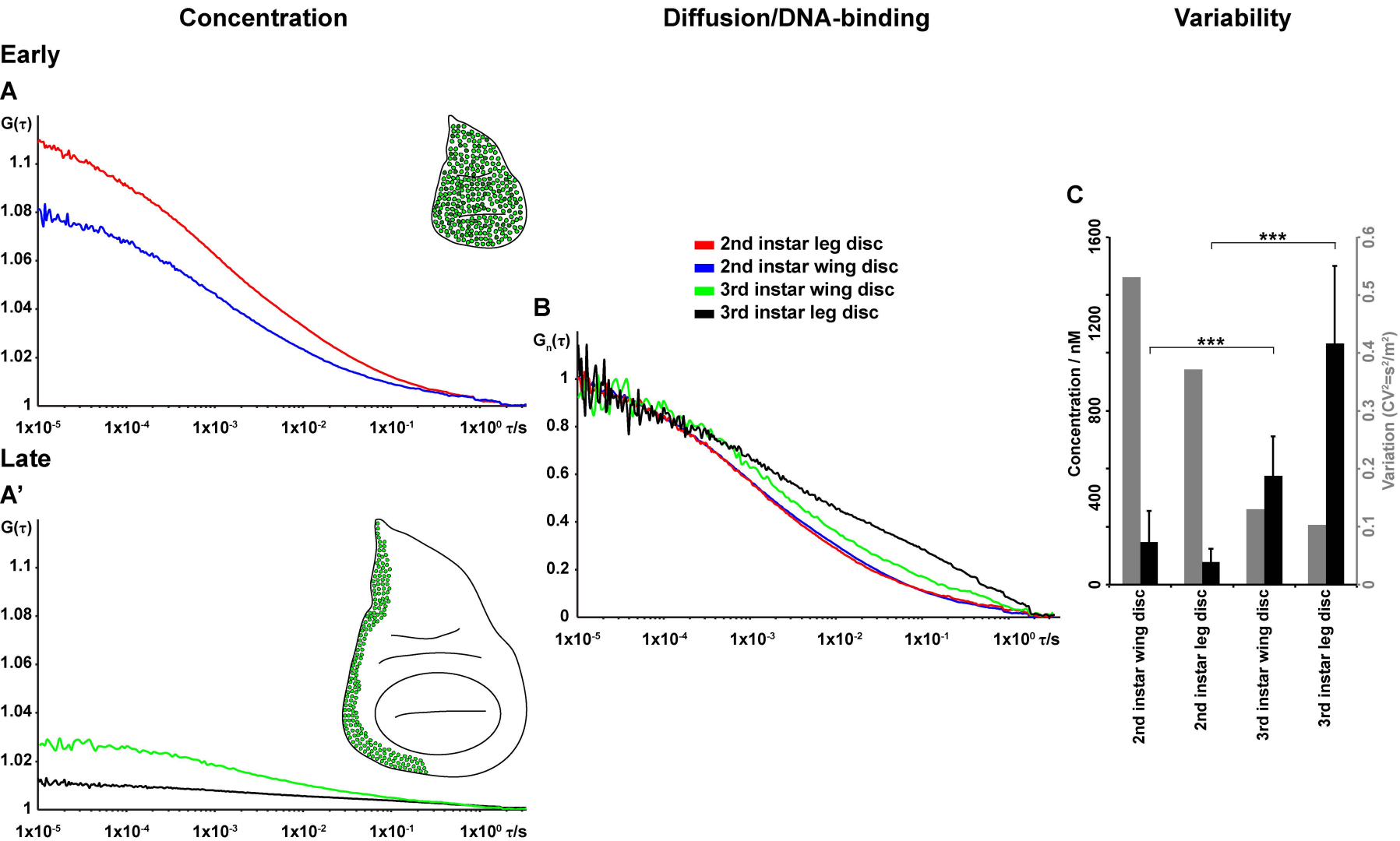
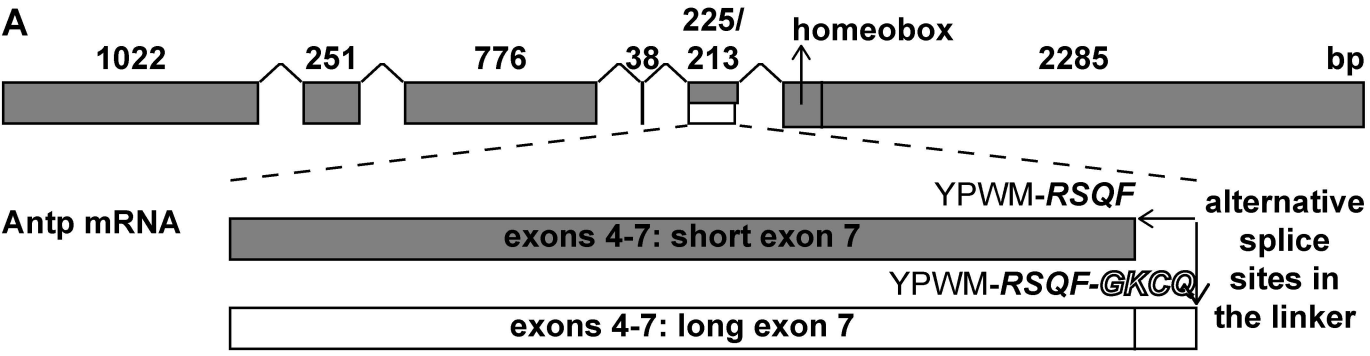
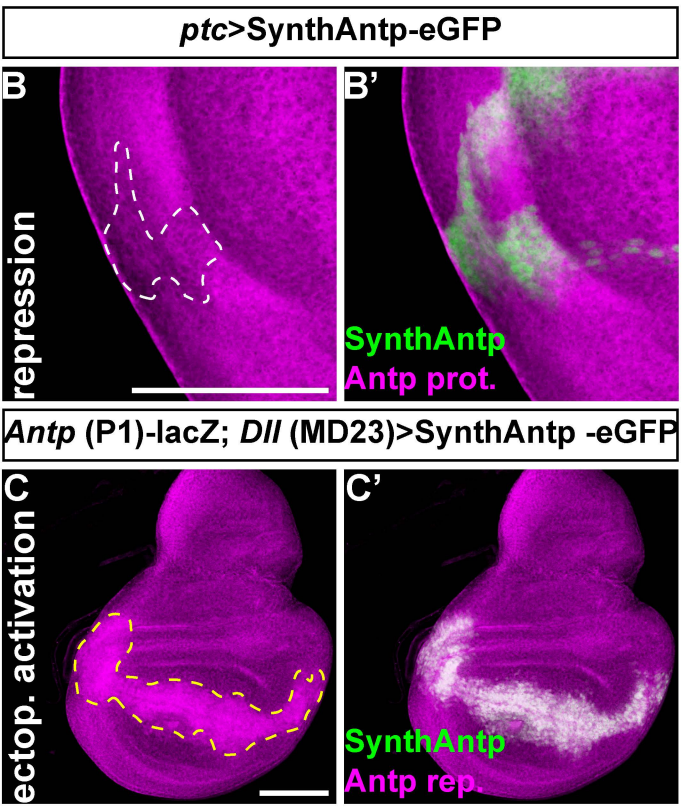
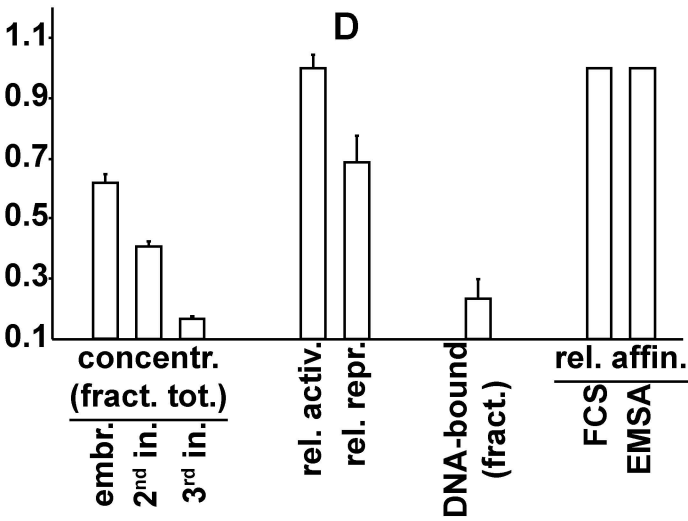


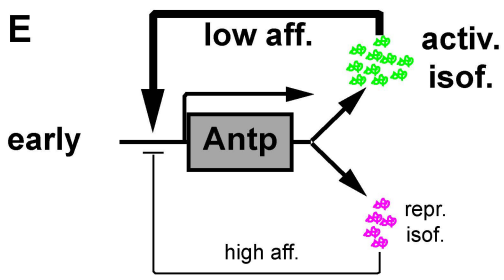
Figure 4



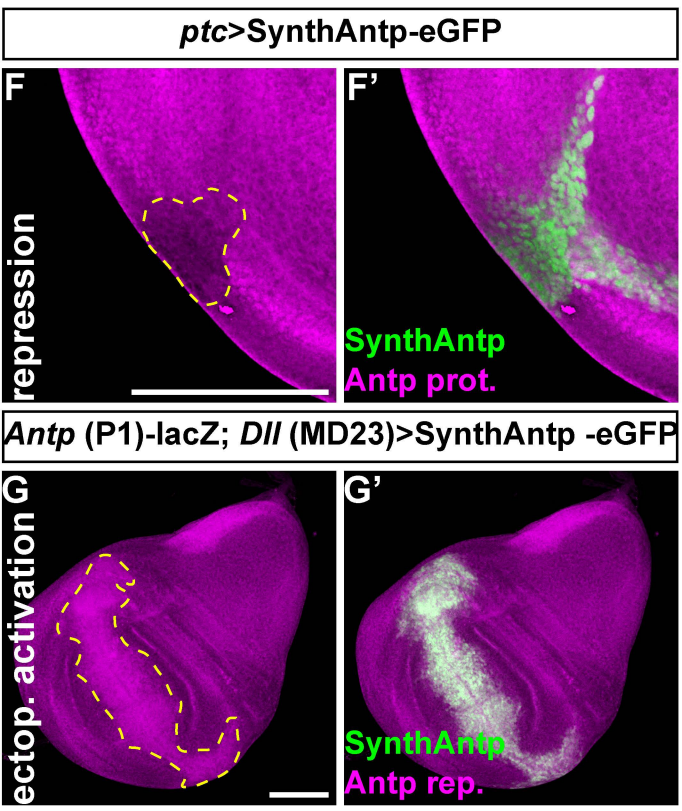
Long linker isoform (preferential activator)

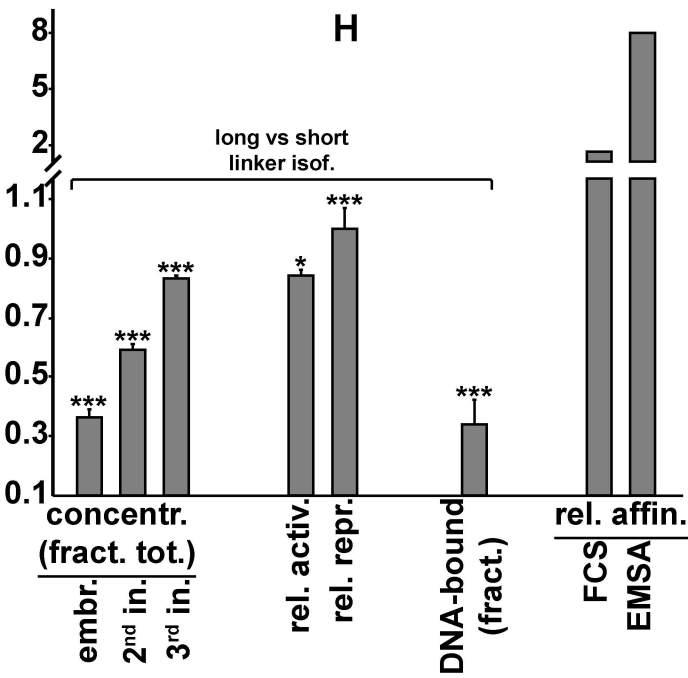






Short linker isoform (preferential repressor)





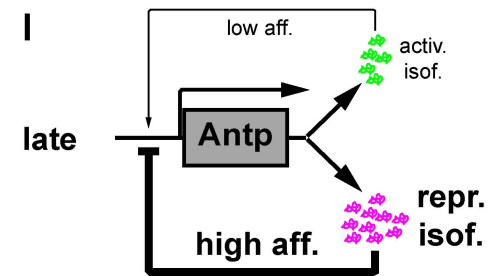


Figure 4 - figure supplement 1

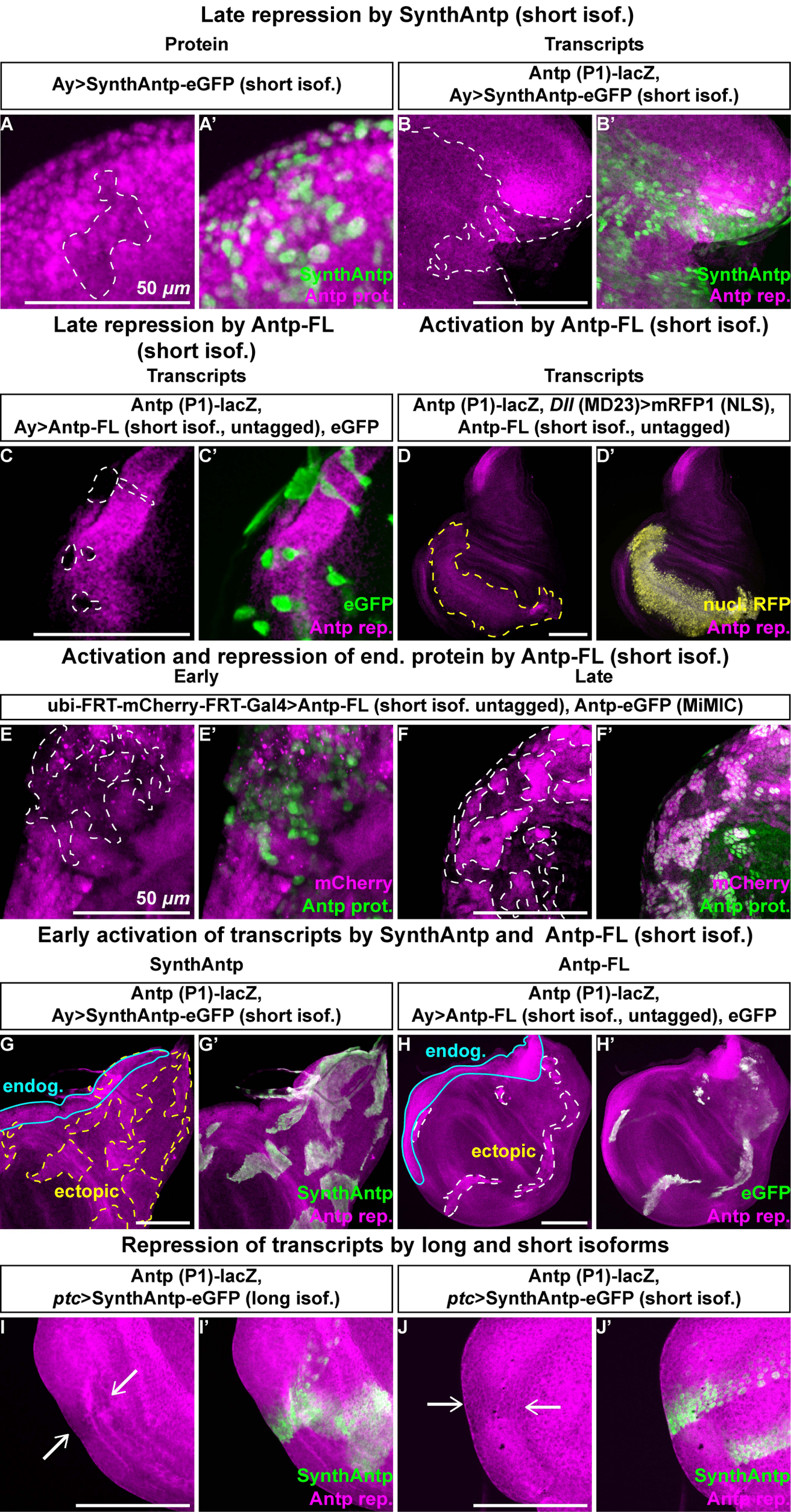


Figure 4 - figure supplement 2

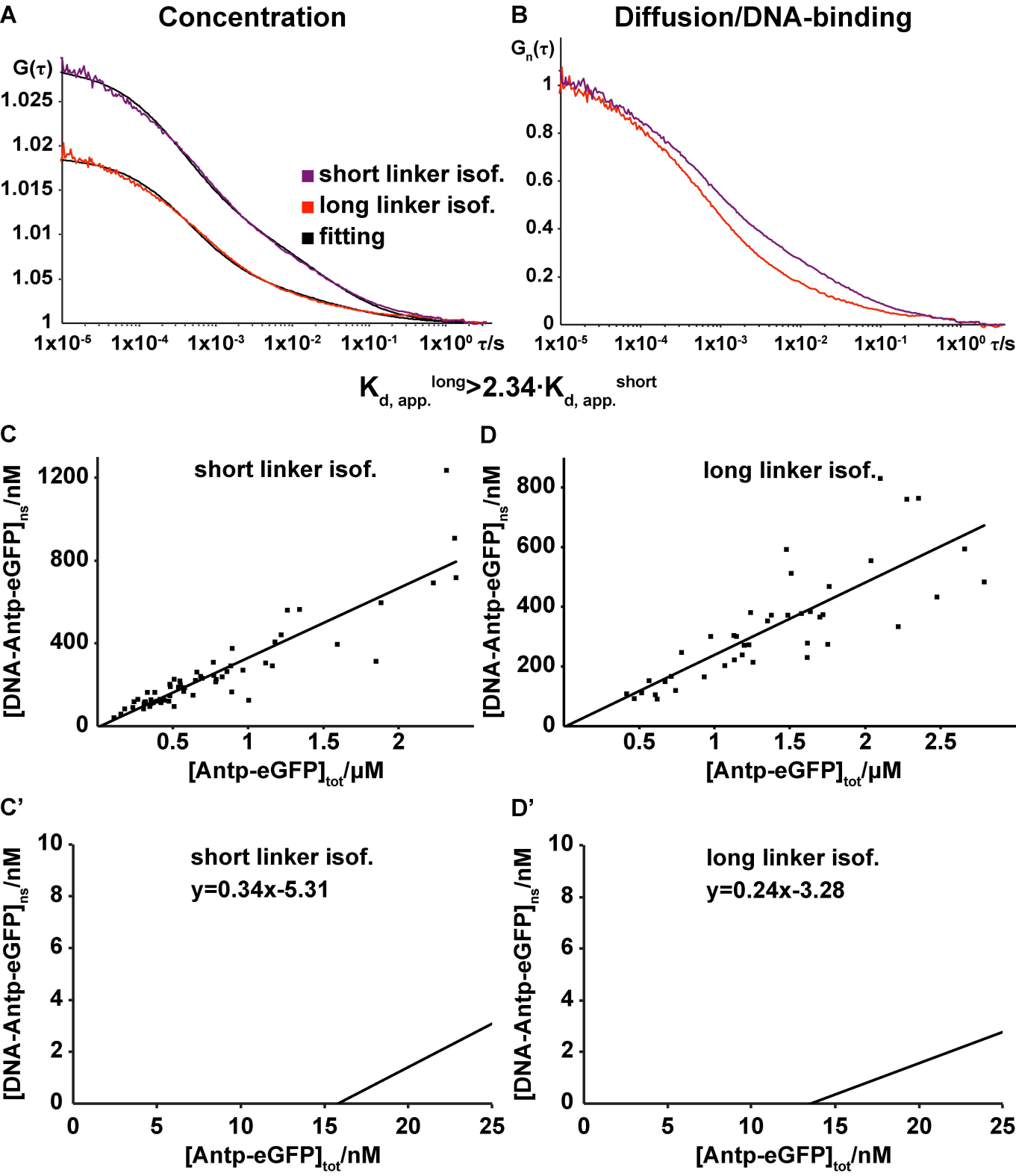


Figure 4 - figure supplement 3

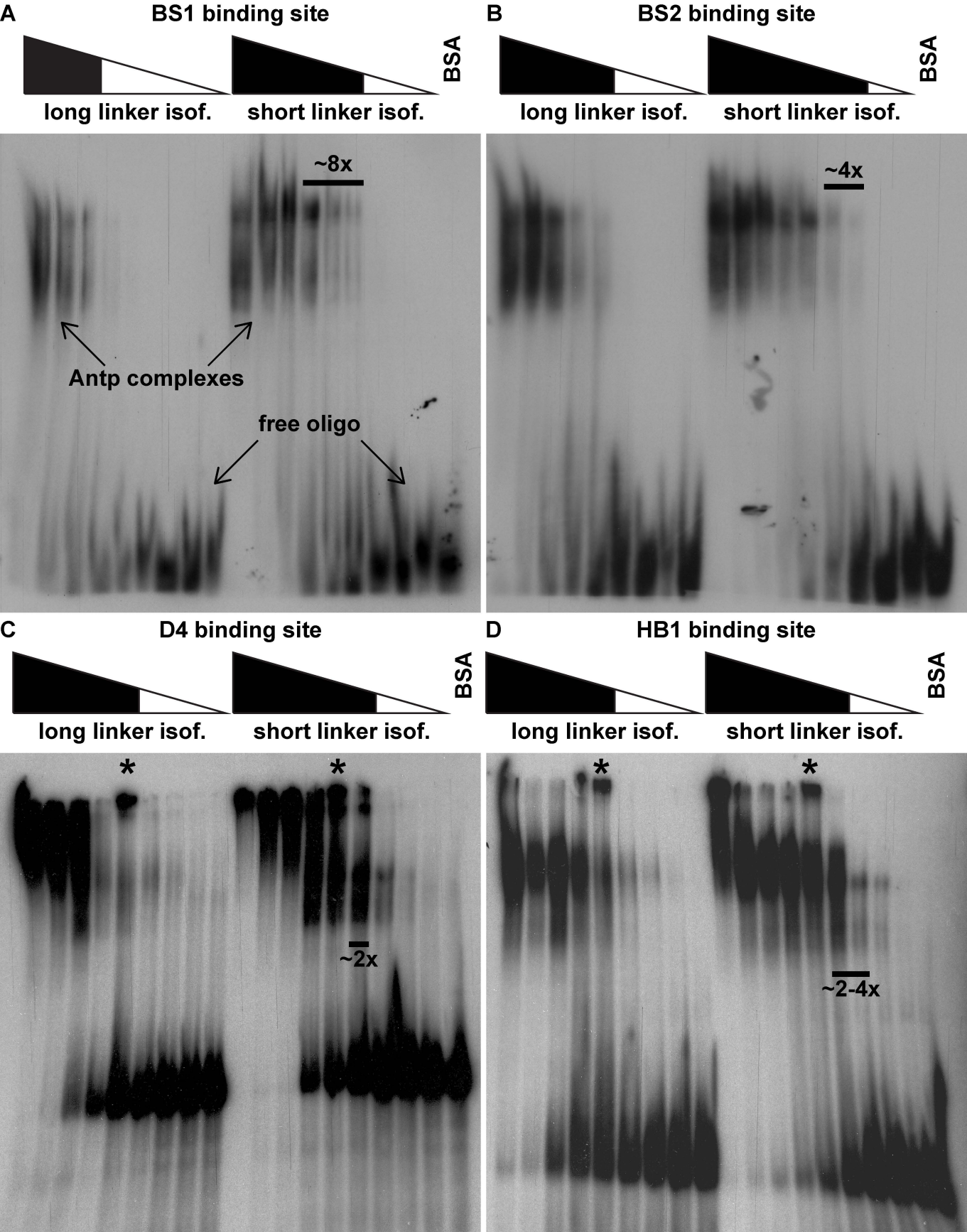


Figure 5

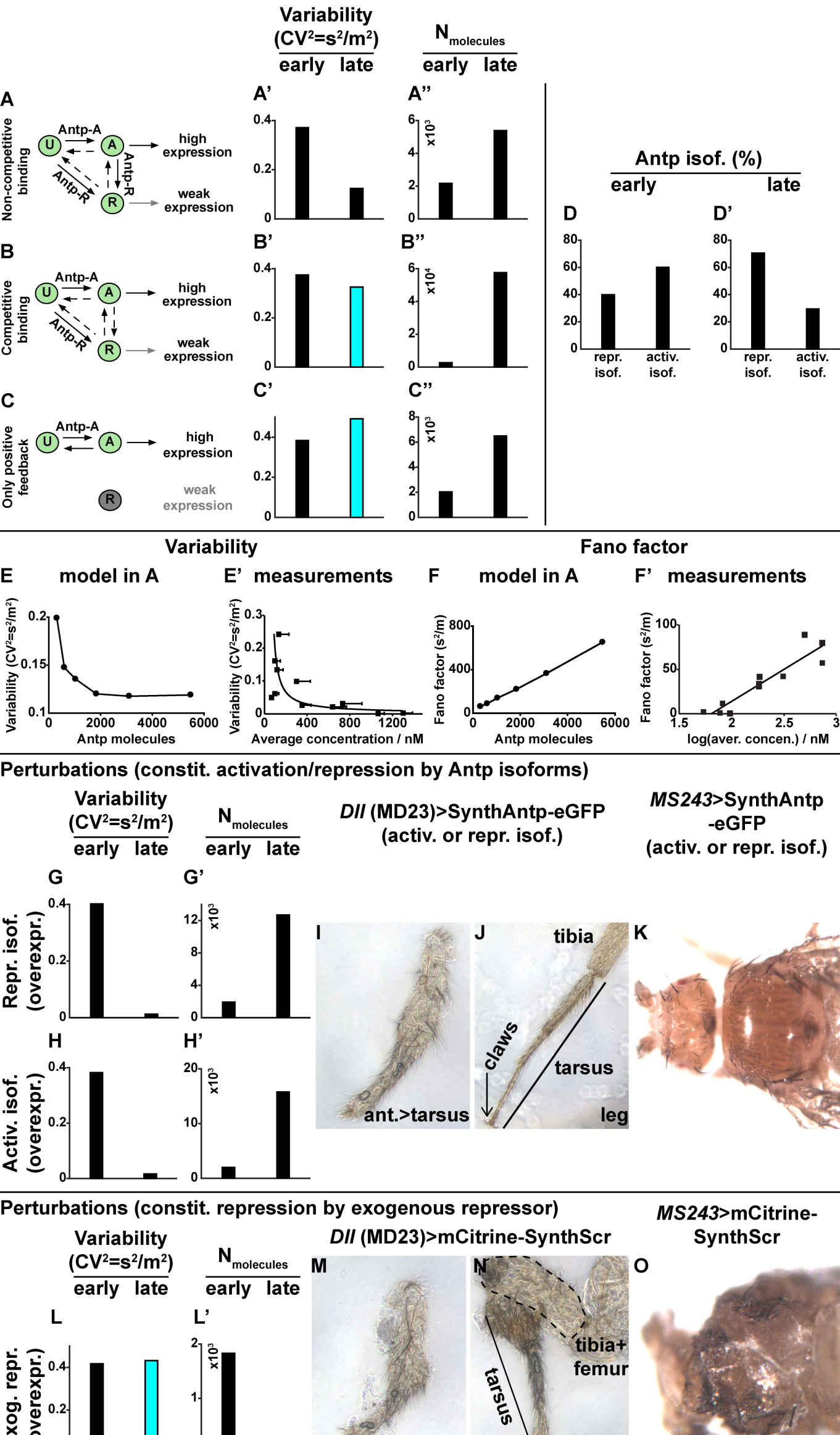

Variability (CV²=s²/m²)
early late
N molecules
early late
A
Non-competitive binding
Antp-A
Antp-R
high expression
weak expression
A'
A''
x10³
B
Competitive binding
B'
B''
x10⁴
C
Only positive feedback
C'
C''
Antp isof. (%)
early late
D
D'
repr. isof.
activ. isof.
Variability
Fano factor
E model in A
E' measurements
F model in A
F' measurements
Variability (CV²=s²/m²)
Antp molecules
Average concentration / nM
Fano factor (s²/m)
log(aver. concen.) / nM
Perturbations (constit. activation/repression by Antp isoforms)
Dll (MD23)>SynthAntp-eGFP (activ. or repr. isof.)
MS243>SynthAntp -eGFP (activ. or repr. isof.)
G
G'
Repr. isof. (overexpr.)
H
H'
Activ. isof. (overexpr.)
I
ant.>tarsus
J
tibia
claws
tarsus
leg
K
Perturbations (constit. repression by exogenous repressor)
Dll (MD23)>mCitrine-SynthScr
MS243>mCitrine-SynthScr
L
L'
Exog. repr. (overexpr.)
M
ant.>tarsus
N
tibia+ femur
tarsus
leg
O
notum

Figure 5 - figure supplement 1

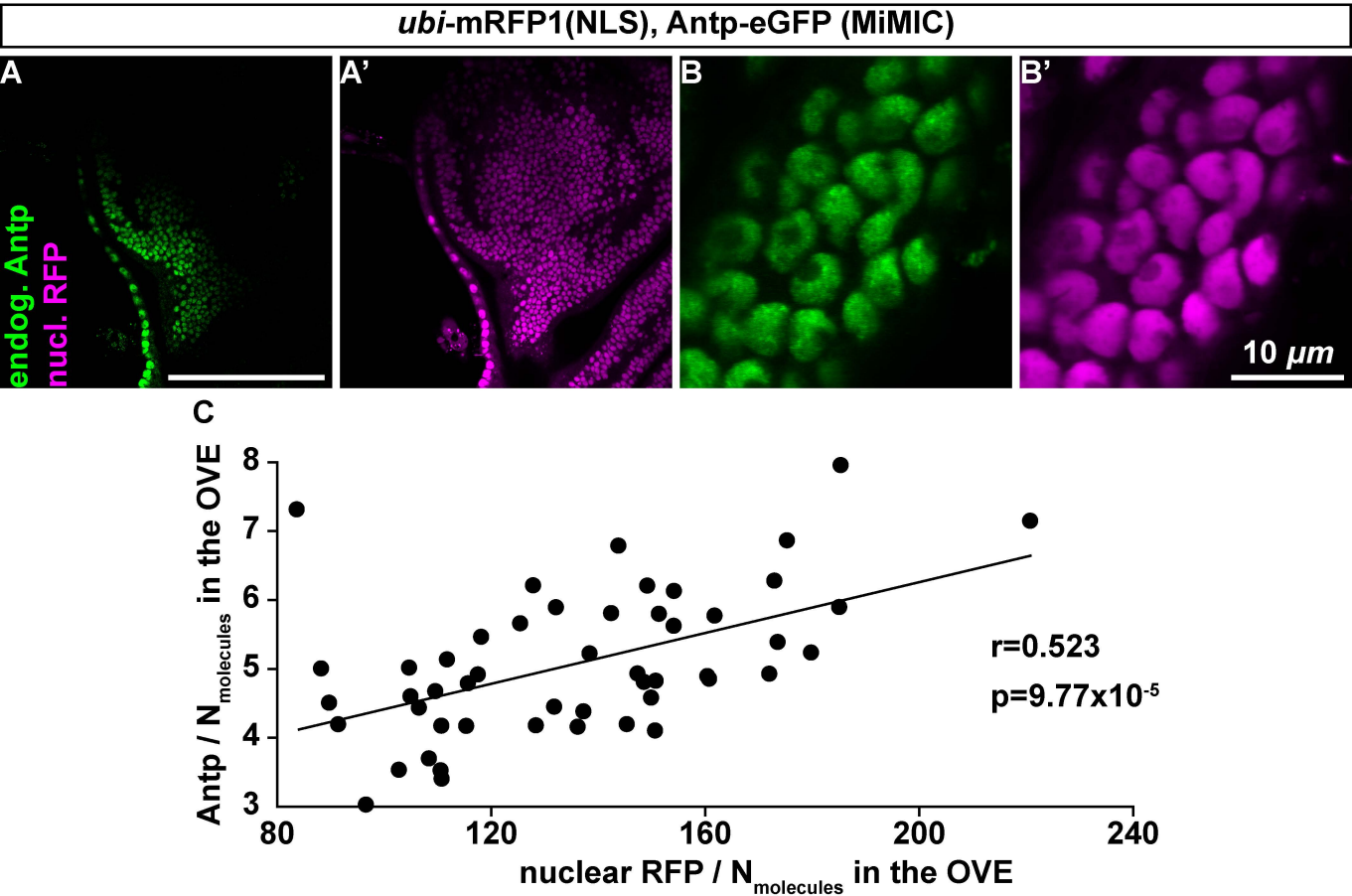
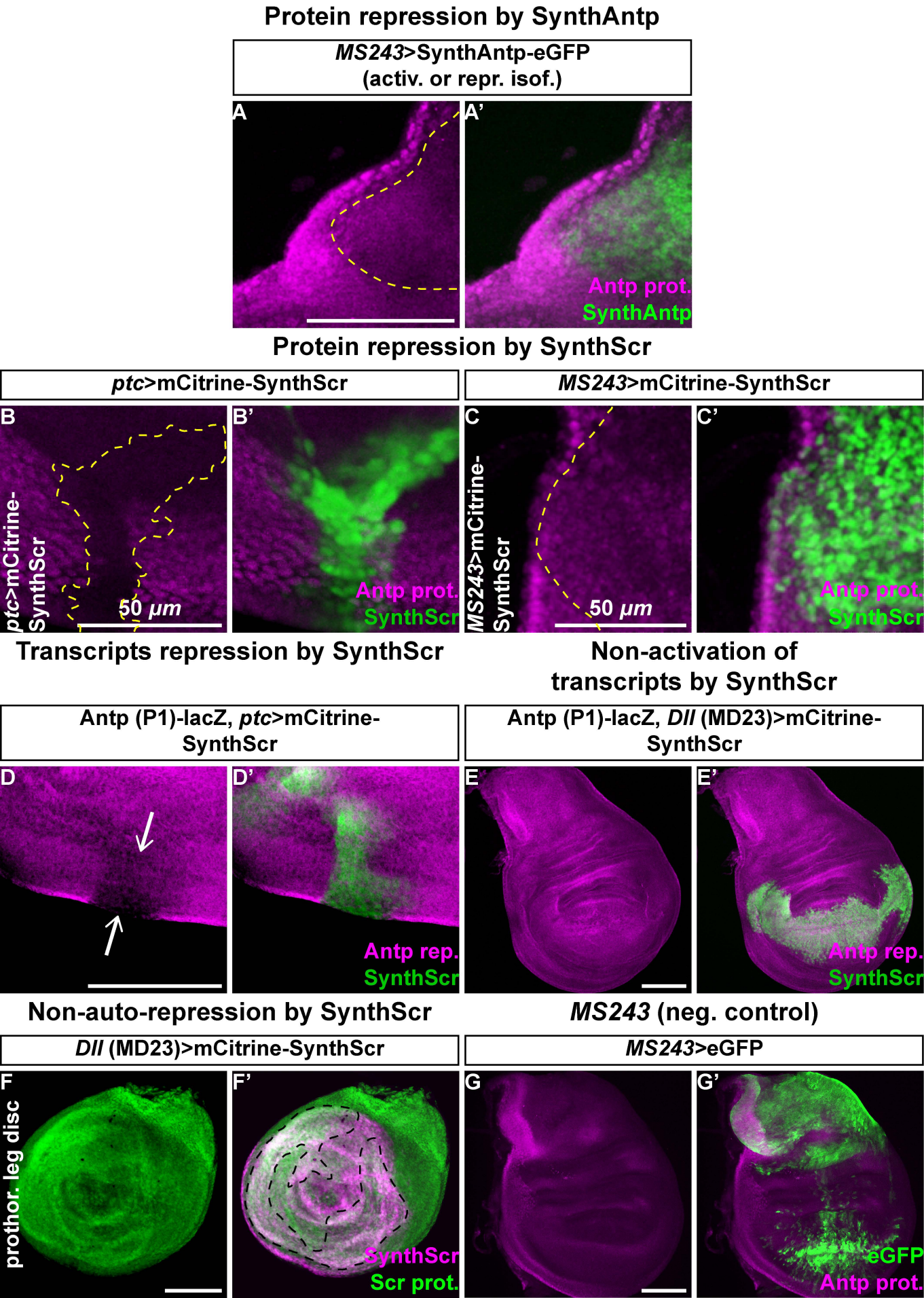



Figure 5 - figure supplement 2